Math 135B, Winter 2023.

**Homework 1**

**Due:** Jan. 20, 2023

1. (a) Assume that $X_n$ are random variables such that that $X_n \to a$ in probability, where $a$ is a (nonrandom) constant. Suppose also $f : \mathbb{R} \to \mathbb{R}$ is a continuous function. Show that $f(X_n) \to f(a)$ in probability.

(b) Assume that $X_n$ and $Y_n$ are random variables such that $X_n$ and $Y_n$ are defined on the same probability space (so that they can be added). Assume also that $X_n \to a$ and $Y_n \to b$ in probability, where $a$ and $b$ are (nonrandom) constants. Show that $X_n + Y_n \to a + b$ in probability.

2. Assume that $X_1, X_2, \dots$ are independent random variables, all uniform on $[0, 1]$. Compute the limit, in probability, of the following random variables:

(a) $\frac{1}{n}\sum_{i=1}^{n} X_n$;

(b) $\frac{1}{n}\sum_{i=1}^{n} X_n^2$;

(c) $\frac{1}{n}\sum_{i=1}^{n} X_i X_{i+1}$; and

(d) $(X_1 \cdots X_n)^{1/n}$.

3. Assume you have $2n$ cards with $n$ colors, with 2 cards of each color. Select $n$ cards without replacement, and let $N_n$ be the number of colors that *are not* represented in your selection.

(a) Compute $EN_n$ and $\mathrm{Var}(N_n)$.

(b) Determine a constant $c$ so that $\frac{1}{n}N_n \to c$, in probability.

(c) Let $M_n$ be the the number of colors that *are* represented in your selection. Determine a constant $d$ so that $\frac{1}{n}M_n \to d$, in probability.

4. (From a Final Exam at Queen's University, Ontario.) An urn contains $m$ red and $n$ blue balls. Balls are drawn one at a time withour replacement until all $m$ red balls are drawn. Let $T$ be the number of draws required. Compute $ET$. (*Hint*. The best way is to relate $T$ to the number $N$ of blue balls that *remain* in the urn after all red balls are drawn.)

Math 135B, Winter 2023.

**Homework 1 Solutions.**

1. (a) Assume that $X_n$ are random variables such that that $X_n \to a$ in probability, where $a$ is a (nonrandom) constant. Suppose also $f : \mathbb{R} \to \mathbb{R}$ is a continuous function. Show that $f(X_n) \to f(a)$ in probability.

Solution. Pick an $\epsilon > 0$. Then there exists a $\delta > 0$ so that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$. This implies the following inclusion of events:

$$\{|f(X_n) - f(a)| \geq \epsilon\} \subset \{|X_n - a| \geq \delta\}.$$

By the assumption, $P(|X_n - a| \geq \delta) \to 0$. Therefore,

$$P(|f(X_n) - f(a)| \geq \epsilon) \leq P(|X_n - a| \geq \delta) \to 0.$$

(b) Assume that $X_n$ and $Y_n$ are random variables such that $X_n$ and $Y_n$ are defined on the same probability space (so that they can be added). Assume also that $X_n \to a$ and $Y_n \to b$ in probability, where $a$ and $b$ are (nonrandom) constants. Show that $X_n + Y_n \to a + b$ in probability.

Solution. Pick an $\epsilon > 0$. By the assumption, $P(|X_n - a| \geq \epsilon/2) \to 0$, $P(|Y_n - b| \geq \epsilon/2) \to 0$. We have

$$|(X_n + Y_n) - (a + b)| = |(X_n - a) + (Y_n - b)| \leq |X_n - a| + |Y_n - b|$$

and so, as events,

$$\{|(X_n + Y_n) - (a + b)| \geq \epsilon\} \subset \{|X_n - a)| \geq \epsilon/2\} \cup \{|Y_n - b)| \geq \epsilon/2\}.$$

It follows that

$$P(|(X_n + Y_n) - (a + b)| \geq \epsilon) \leq P(|X_n - a)| \geq \epsilon/2) + P(|Y_n - b)| \geq \epsilon/2) \to 0.$$

2. Assume that $X_1, X_2, \dots$ are independent random variables, all uniform on $[0, 1]$. Compute the limit, in probability, of the following random variables:
(a) $\frac{1}{n}\sum_{i=1}^{n} X_n$;

Solution. By WLLN, the limit is $EX_1 = 1/2$.

(b) $\frac{1}{n}\sum_{i=1}^{n} X_n^2$;

Solution. By WLLN, the limit is $EX_1^2 = 1/3$.

(c) $\frac{1}{n}\sum_{i=1}^{n} X_i X_{i+1}$; and

Solution.
Let $S_n$ be the sum. We need to compute the limit of $S_n/n$.

*First solution*. We use the second moment method. By independence, $E(X_iX_{i+1}) = EX_iEX_{i+1} = 1/4$, and

$$\mathrm{Var}(X_iX_{i+1}) = E(X_i^2X_{i+1}^2) - \left(\frac{1}{4}\right)^2 = \frac{1}{9} - \frac{1}{16} = \frac{7}{144}.$$

Also, assume $j > i$. If $j \geq i+2$, then

$$\mathrm{Cov}(X_iX_{i+1}, X_jX_{j+1})) = 0,$$

again by independence. If $j = i+1$, then

$$\mathrm{Cov}(X_iX_{i+1}, X_{i+1}X_{i+2})) = EX_iEX_{i+1}^2EX_{i+2} - \frac{1}{16} = \frac{1}{2\cdot 3\cdot 2} - \frac{1}{16} = \frac{1}{48}.$$

(In fact, the precise values of the variance and the covariance when $j = i+1$ are not important, as long as they are finite. What is important is that the covariance vanishes when $j > i+1$. There are many problems with "local dependence" in which a similar property holds.) We conclude that, by the variance-covariance formula,

$$\begin{aligned}\mathrm{Var}(S_n) &= \sum_{i=1}^n \mathrm{Var}(X_iX_{i+1}) + 2\sum_{i<j}\mathrm{Cov}(X_iX_{i+1}, X_jX_{j+1})\\ &= \sum_{i=1}^n \mathrm{Var}(X_iX_{i+1}) + 2\sum_{i=1}^{n-1}\mathrm{Cov}(X_iX_{i+1}, X_{i+1}X_{i+2})\\ &= n\frac{7}{144} + 2(n-1)\frac{1}{48}.\end{aligned}$$

It follows that $E(S_n/n) = 1/4$ and $\mathrm{Var}(S_n/n) = \mathrm{Var}(S_n)/n^2 = \mathcal{O}(1/n))$, so that $S_n/n \to 1/4$ in probability.

*Second solution*. Assume first that $n$ is odd. Write

$$\frac{1}{n}S_n = \frac{1}{n}(X_1X_2 + X_3X_4 + \cdots X_nX_{n+1}) + \frac{1}{n}(X_2X_3 + X_4X_5 + \cdots X_{n-1}X_n)$$

Call the first sum $S_n'$ and the second $S_n''$. Then $S_n'$ has $(n+1)/2$ independent (and identically distributed) terms, each with with expectation $1/4$, and $S_n''$ has $(n-1)/2$ independent terms, each with with expectation $1/4$. So, $\frac{2}{n+1}S_n' \to \frac{1}{4}$ and $\frac{2}{n-1}S_n'' \to \frac{1}{4}$, in probability, by WLLN. Our sum is

$$\frac{1}{n}S_1 + \frac{1}{n}S_2 = \frac{n+1}{2n}\cdot\frac{2}{n+1}S_1 + \frac{n-1}{2n}\cdot\frac{2}{n-1}S_2 \to \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{4},$$

by Problem 1(b). The case when $n$ is even is similar. The answer is $1/4$.

(d) $(X_1\cdots X_n)^{1/n}$.

<u>Solution</u>. Rewrite as $\exp(\frac{1}{n}\sum_{i=1}^n \log X_n)$. By WLLN,

$$\frac{1}{n}\sum_{i=1}^n \log X_n \to E\log X_1 = \int_0^1 \log x\,dx = -1,$$

in probability, and so the limit is $e^{-1}$, by problem 1(a), applied to the function $f(x) = e^x$.

*Note*. To apply WLLN as we stated in class, we also need to verify that $EX_1^2 = \int_0^1 (\log x)^2\,dx$ is finite. (In fact, WLLN holds without that assumption.) By the change of variables $\log x = -z$, we get $EX_1^2 = \int_0^\infty z^2 e^{-z}\,dz = 2$.

3. Assume you have $2n$ cards with $n$ colors, with 2 cards of each color. Select $n$ cards without replacement, and let $N_n$ be the number of colors that *are not* represented in your selection.

(a) Compute $EN_n$ and $\text{Var}(N_n)$.

Solution. Let $I_i$ be the indicator of the event that color $i$ is missing. Then $N_n = I_1 + \ldots + I_n$. The indicators all have the same expectation

$$EI_i = P(\text{color } i \text{ is missing}) = \frac{\binom{2n-2}{n}}{\binom{2n}{n}} = \frac{(2n-2)!n!}{(2n)!(n-2)!} = \frac{n-1}{2(2n-1)},$$

and so

$$EN_n = \frac{n(n-1)}{2(2n-1)}.$$

Moreover, for $i \neq j$,

$$E(I_iI_j) = P(\text{colors } i,j \text{ are both missing}) = \frac{\binom{2n-4}{n}}{\binom{2n}{n}} = \frac{(2n-4)!n!}{(2n)!(n-4)!}$$
$$= \frac{n(n-1)(n-2)(n-3)}{2n(2n-1)(2n-2)(2n-3)} = \frac{(n-2)(n-3)}{4(2n-1)(2n-3)}.$$

It follows that

$$\text{Var}(N_n) = EN_n + \sum_{i\neq j} E(I_iI_j) - EN_n^2$$
$$= \frac{n(n-1)}{2(2n-1)} + n(n-1)\frac{(n-2)(n-3)}{4(2n-1)(2n-3)} - \left(\frac{n(n-1)}{2(2n-1)}\right)^2.$$

(It is not necessary to simplify further.)

(b) Determine a constant $c$ so that $\frac{1}{n}N_n \to c$, in probability.

Solution. Let $X_n = N_n/n$. By (a),

$$EX_n = \frac{EN_n}{n} \to \frac{1}{4},$$

and

$$\text{Var}(X_n) = \frac{1}{n^2}\text{Var}(N_n) \to 0 + \frac{1}{16} - \frac{1}{16} = 0.$$

Therefore, the convergence holds with $c = 1/4$.

(c) Let $M_n$ be the the number of colors that *are* represented in your selection. Determine a constant $d$ so that $\frac{1}{n}M_n \to d$, in probability.

Solution. As $M_n = n - N_n$, $\frac{1}{n}M_n = 1 - \frac{1}{n}N_n$, and we may take $d = 1 - c = 3/4$.

4. (From a Final Exam at Queen's University, Ontario.) An urn contains $m$ red and $n$ blue balls. Balls are drawn one at a time withour replacement until all $m$ red balls are drawn. Let $T$ be the number of draws required. Compute $ET$. (*Hint*. The best way is to relate $T$ to the number $N$ of blue balls that *remain* in the urn after all red balls are drawn.)

<u>Solution</u>. Following the hint, we observe that $T = m + n - N$. To each blue ball $i$, we attach the indicator $I_i$ of the event that it remains in the urn after all red balls are selected. This is the event that the ball $i$ is the last in the ordering of $m$ red balls and the blue ball $i$, and thus has the probability $1/(m+1)$. So $EI_i = 1/(m+1)$. Furthermore, $N = I_1 + \ldots + I_n$, so $EN = n/(m+1)$ and

$$ET = m + n - \frac{n}{m+1} = \frac{m(m+n+1)}{m+1}.$$