

PROBLEMS AND SOLUTIONS:  
*Theory of Probability and Random Processes, 2<sup>nd</sup> Edition*

Leonid B. Korolov and Yakov G. Sinai

Alan Sun

# Random Variables and Their Distributions

## Problem 1.

*Solution.*  $300 \times 10^6 \times 0.5 \times \frac{2}{365^3}$

## Problem 2.

*Solution.* Suppose  $n$  identical balls are thrown into  $m$  boxes. Let  $\Omega$  be the set of  $m$ -tuples where each entry is a non-negative integer and the sum of the entries is  $n$ . Then, by a stars and bars argument, we know that  $|\Omega| = \binom{n+m-1}{m-1}$ . Therefore,

$$\mathbb{P}[\text{first box is empty}] = \frac{\binom{n+m-2}{m-2}}{\binom{n+m-1}{m-1}} = \frac{m-1}{n+m-1}.$$

## Problem 3.

*Solution.* Let  $\Omega$  be the set of tuples where the first entry is the number of defective items and the second is the number of good items. Then,  $\mathbb{P}[(0, 10)] = \frac{\binom{90}{10}}{\binom{100}{100}} = \frac{90 \cdot 89 \cdot \dots \cdot 81}{100 \cdot 99 \cdot \dots \cdot 91}$ .

## Problem 4.

*Solution.* By way of contradiction, suppose that  $\mathbb{P}[|\xi| > C] > 0$  for  $C > 0$ . Then, it follows that there exists some  $A' \geq 1$  such that  $\mathbb{P}[|\xi| \geq A'C] > 0$ . So,

$$\begin{aligned} \mathbb{E}[|\xi|^m] &= \sum_{\omega \in \Omega} |\xi(\omega)|^m, \\ &\geq \sum_{\omega \in \{|\xi| \geq A'C\}} |\xi(\omega)|^m, \\ &\geq A'^m C^m \mathbb{P}[|\xi| \geq A'C]. \end{aligned}$$

Since  $\mathbb{P}[|\xi| \geq A'C]$  is constant, clearly there does not exist  $A > 0$  such that  $A \geq A'^m$  for all  $m$ , thus a contradiction.

## Problem 5.

*Solution.* This is the famous hat-check problem. We find the probability that no one gets the right letter. Let  $\Omega = S_n$ , the permutation group of  $n$  elements, also let  $C_i$  be the set of all permutations where the  $i^{\text{th}}$  element is in the  $i^{\text{th}}$  position. Formally,  $\sigma \in C_i$  if and only if  $\sigma(i) = i$ . Then, by the inclusion-exclusion principle, we have that

$$\begin{aligned}\mathbb{P}[\{\sigma : \exists i, \sigma(i) = i\}] &= \mathbb{P}\left[\bigcup_{i=1}^n C_i\right], \\ &= \sum_{i=1}^n \mathbb{P}[C_i] - \sum_{i < j} \mathbb{P}[C_i \cap C_j] + \sum_{i < j < k} \mathbb{P}[C_i \cap C_j \cap C_k] - \dots\end{aligned}$$

Now, we derive the expression for each term in the previous sum. For any  $\sigma \in C_{i_1} \cap C_{i_2} \cap \dots \cap C_{i_k}$ ,  $\sigma(i_1) = i_1$ ,  $\sigma(i_2) = i_2, \dots, \sigma(i_k) = i_k$ . Thus,  $\mathbb{P}[C_{i_1} \cap C_{i_2} \cap \dots \cap C_{i_k}] = \frac{(n-k)!}{n!}$ . Therefore,  $\sum_{i_1 < i_2 < \dots < i_k} \mathbb{P}[C_{i_1} \cap C_{i_2} \cap \dots \cap C_{i_k}] = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}$ . So,

$$\begin{aligned}\mathbb{P}[\{\sigma : \exists i, \sigma(i) = i\}] &= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots + \frac{(-1)^{n+1}}{n!}, \\ \lim_{n \rightarrow \infty} \mathbb{P}[\{\sigma : \exists i, \sigma(i) = i\}] &= \frac{1}{e}.\end{aligned}$$

### Problem 6.

*Solution.* The first part of the question is a direct application of a stars and bars argument. Then, there are  $\binom{n+r-1}{r-1}$  solutions to the equation  $x_1 + \dots + x_r = n$ . By the same argument, we find that

$$\mathbb{P}[x_1 = a] = \frac{\binom{n-a+r-2}{r-2}}{\binom{n+r-1}{r-1}} = \frac{(r-1)n!(n-a+r-2)!}{(n+r-1)!(n-a)!}.$$

Now, we take the limit as  $r, n \rightarrow \infty$  and  $n/r \rightarrow \rho > 0$ .

$$\begin{aligned}\lim_{\substack{n, r \rightarrow \infty \\ n/r \rightarrow \rho}} \frac{(r-1)n!(n-a+r-2)!}{(n+r-1)!(n-a)!} &= \lim_{\substack{n, r \rightarrow \infty \\ n/r \rightarrow \rho}} \frac{(r-1)n(n-1) \dots (n-a+1)}{(n+r-1)(n+r-2) \dots (n-a+r-1)}, \\ &\approx \lim_{\substack{n, r \rightarrow \infty \\ n/r \rightarrow \rho}} \frac{rn^a}{(n+r)^{a+1}}, \\ &= \frac{r(\rho r)^a}{(\rho r + r)^{a+1}}, \\ &= \frac{\rho^a}{(\rho + 1)^a}.\end{aligned}$$

### Problem 7.

*Solution.* Recall, that the Poisson distribution is the measure on  $\mathbb{Z}^+$  such that for any elementary outcome  $k$ ,  $\mathbb{P}[k] = \frac{\lambda^k}{e^\lambda k!}$ . Let  $\xi = \text{Id}_{\mathbb{Z}^+}$ . Then,

$$\begin{aligned}\mathbb{E}[\xi] &= \sum_{n=0}^{\infty} \xi(n) \mathbb{P}[n], \\ &= \sum_{n=1}^{\infty} \frac{\lambda^n}{e^\lambda (n-1)!},\end{aligned}$$

expanding the sum to see the Taylor expansion more clearly, we see that

$$\begin{aligned}
&= \frac{\lambda}{e^\lambda} \left( 1 + \lambda + \frac{\lambda^2}{2!} + \dots \right), \\
&= \frac{\lambda}{e^\lambda} e^\lambda, \\
&= \lambda.
\end{aligned}
\tag{by Taylor expansion}$$

Now, we find the variance using the formula  $\text{Var}[\xi] = \mathbb{E}[\xi^2] - (\mathbb{E}[\xi])^2$ . So,

$$\begin{aligned}
\mathbb{E}[\xi^2] &= \sum_{n=0}^{\infty} \xi(n)^2 \mathbb{P}[n], \\
&= e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n n}{(n-1)!}, \\
&= \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n (n+1)}{n!}, \\
&= \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} + \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!}, \\
&= \lambda e^{-\lambda} \left( e^\lambda + \lambda \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \right), \\
&= \lambda e^{-\lambda} (e^\lambda + \lambda e^\lambda), \\
&= \lambda^2 + \lambda.
\end{aligned}$$

Therefore,  $\text{Var}[\xi] = \mathbb{E}[\xi^2] - (\mathbb{E}[\xi])^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$  and the variance of the Poisson distribution is the same as its expected value.

### Problem 9.

*Solution.* Recall that if  $F$  is a distribution function of a random variable  $\xi$  then,  $F_\xi(x) = \mathbb{P}[\{\xi(\omega) < x : \omega \in \Omega\}]$ . We then evaluate the expression on the right-hand side first

$$\mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq x\}] - \lim_{\delta \downarrow 0} \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq x - \delta\}] = \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq x\}] - \mathbb{P}\left[\bigcup_{i=1}^{\infty} \{\omega \in \Omega : \xi(\omega) \leq x - \delta_i\}\right],$$

where  $\delta_1, \delta_2, \dots$  is any sequence where  $\delta_i \geq 0$  and converges to 0. We first show that

$$\bigcup_{i=1}^{\infty} \{\omega \in \Omega : \xi(\omega) \leq x - \delta_i\} = \{\omega \in \Omega : \xi(\omega) < x\},$$

That the left-hand side is a subset of the right-hand side is trivial. Now, suppose an element  $\omega$  in the set on the right-hand side. There must exist some  $\varepsilon > 0$ , such that  $\xi(\omega) = x - \varepsilon$ . By convergence of  $\delta_1, \delta_2, \dots$ , there must exist some  $\delta_i$  where  $\xi(\omega) = x - \varepsilon \leq x - \delta_i$ . Thus, the two sets are equivalent. Therefore,

$$\begin{aligned}
\mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq x\}] - \mathbb{P}\left[\bigcup_{i=1}^{\infty} \{\omega \in \Omega : \xi(\omega) \leq x - \delta_i\}\right] &= \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq x\}] - \mathbb{P}[\{\omega \in \Omega : \xi(\omega) < x\}], \\
&= \mathbb{P}[\{\omega \in \Omega : \xi(\omega) = x\}], \\
&= \mathbb{P}[\xi = x].
\end{aligned}$$

The second equality follows  $\sigma$ -additivity of  $\mathbb{P}$ .

### Problem 10.

*Solution.* Let  $F$  be the distribution of  $\xi$  and let  $\eta = a\xi + b$  where  $a \neq 0$ . We first find the distribution of  $\eta$ :

$$\begin{aligned}\mathbb{P}[\{\omega \in \Omega : \eta(\omega) \leq x\}] &= \mathbb{P}[\{\omega \in \Omega : a\xi(\omega) + b \leq x\}], \\ &= \mathbb{P}\left[\left\{\omega \in \Omega : \xi(\omega) \leq \frac{x-b}{a}\right\}\right], \\ &= F\left(\frac{x-b}{a}\right).\end{aligned}$$

Thus, taking the derivative to yield the density function, we have that  $p_\eta(x) = \frac{1}{a}p_\xi\left(\frac{x-b}{a}\right)$  where  $p_\xi$  is the density of  $\xi$ .

### Problem 11.

*Solution.* To show this we need a few simple measure-theoretic facts:

1. By definition, the density function of a random variable uniformly distributed on some set  $A \subset \mathbb{R}$  where  $\lambda(A) < \infty$  is  $\frac{1}{\lambda(A)}$ . Moreover, for any measurable-subset  $B \subset A$ ,  $\mathbb{P}_A[B] = \frac{\lambda(B)}{\lambda(A)}$ . For brevity, We call  $\mathbb{P}_A$  the Lebesgue probability measure on  $A$ .
2. Let  $C \subset D \subset G \subset \mathbb{R}$  (be measurable-subsets of  $\mathbb{R}$ ) and  $\mathbb{P}_D, \mathbb{P}_G$  be the Lebesgue probability measures on  $D, G$ , respectively. Then, by definition,  $\mathbb{P}_D[C] = \frac{\lambda(C)}{\lambda(D)}$  and  $\mathbb{P}_G[D] = \frac{\lambda(D)}{\lambda(G)}$ . It follows that  $\mathbb{P}_G[C] = \frac{\lambda(C)}{\lambda(G)} = \mathbb{P}_D[C]\mathbb{P}_G[D]$ .

Let  $(\Omega_1, \mathcal{G}_1, \mathbb{P}_{\Omega_1})$  and  $(\Omega_2, \mathcal{G}_2, \mathbb{P}_{\Omega_2})$  be the probability subspaces of  $[0, 2\pi]$  with the Lebesgue probability measure, where  $\Omega_1 = [0, \pi/2] \cup [3\pi/2, 2\pi]$  and  $\Omega_2 = [\pi/2, 3\pi/2]$ . Also, let  $\mathbb{P}_{\Omega_1}, \mathbb{P}_{\Omega_2}$  be the Lebesgue probability measures on  $\Omega_1, \Omega_2$ , respectively. Suppose  $\xi = \text{Id}_{[0, 2\pi]}$ ,  $\xi_1 = \text{Id}_{[0, \pi/2] \cup [3\pi/2, 2\pi]}$ ,  $\xi_2 = \text{Id}_{[\pi/2, 3\pi/2]}$ , and  $\eta = \sin(\xi)$ . Then, we start by finding the distribution of  $\eta$ :

$$\begin{aligned}\mathbb{P}[\{\omega \in [0, 2\pi] : \eta(\omega) \leq x\}] &= \mathbb{P}[\{\omega \in [0, \pi/2] \cup [3\pi/2, 2\pi] : \eta(\xi(\omega)) \leq x\}] \\ &\quad + \mathbb{P}[\{\omega \in [\pi/2, 3\pi/2] : \eta(\xi(\omega)) \leq x\}], \\ &= \frac{1}{2}\mathbb{P}_{\Omega_1}[\{\omega \in \Omega_1 : \eta(\xi_1(\omega)) \leq x\}] + \frac{1}{2}\mathbb{P}_{\Omega_2}[\{\omega \in \Omega_2 : \eta(\xi_2(\omega)) \leq x\}], \\ &= \frac{1}{2}F_1(\arcsin(x)) + \frac{1}{2}F_2(\arcsin(x)),\end{aligned}$$

where  $F_1, F_2$  are the distribution functions for  $\xi_1, \xi_2$ , respectively. Differentiating with respect to  $x$ , we yield the density function:

$$\begin{aligned}\frac{1}{2}(F_1(\arcsin(x)) + F_2(\arcsin(x)))' &= \frac{1}{2\sqrt{1-x^2}}(F_1'(\arcsin(x)) + F_2'(\arcsin(x))), \\ &= \frac{1}{2\sqrt{1-x^2}}\left(\frac{1}{\pi} + \frac{1}{\pi}\right), \\ &= \frac{1}{\pi\sqrt{1-x^2}}.\end{aligned}$$

### Problem 12.

*Solution.* Let  $\mathcal{B}(X)$  be the Borel sets of  $X$ , the  $\sigma$ -algebra of all of the open sets in  $X$ . Also let  $\mathcal{B}(X_1) \times \dots \times \mathcal{B}(X_n)$  be the product  $\sigma$ -algebra of  $X_1, \dots, X_n$ . That  $\mathcal{B}(X_1) \times \dots \times \mathcal{B}(X_n) \subset \mathcal{B}(X)$  is trivial since the product of open sets in  $X_1, \dots, X_n$  must be an open set in  $X$ . It remains to show inclusion in the other direction. Since  $X$  is a separable metric space, it is second-countable. Moreover, we know that  $B = \{x_1 \times \dots \times x_n : x_1 \in X_1, \dots, x_n \in X_n\}$  forms a basis of the product topology. Thus, any open set in  $X$  can be constructed

through the countable unions of sets in  $B$ . And so, since  $\sigma$ -algebras are closed under countable unions the inclusion holds in the other direction.

### Problem 13.

*Solution.* For any  $n > 5$ ,  $n^4 \geq 1000$ . Therefore, we are mostly concerned with the set of all perfect squares and cubes less than or equal to 1000. Thus, we find these first using the inclusion-exclusion principle. Then, brute-force count the remaining numbers that do not fall into these sets.

$$\begin{aligned} \# \{1 \leq n \leq 1000 : n^2 \leq 1000 \text{ or } n^3 \leq 1000\} &= \# \{1 \leq n \leq 1000 : n^2 \leq 1000\} + \# \{1 \leq n \leq 1000 : n^3 \leq 1000\} \\ &\quad - \# \{1 \leq n \leq 1000 : n^6 \leq 1000\}, \\ &= \lfloor \sqrt{1000} \rfloor + \lfloor \sqrt[3]{1000} \rfloor - 2, \\ &= 31 + 10 - 3, \\ &= 39. \end{aligned}$$

It remains to account for numbers where the exponential by 5, 7 are within 1000. This only holds for  $2^5, 2^7, 3^5$ . Therefore, there are 41 numbers less than or equal to 1000 that are integer powers (greater than 1) of another integer. Thus, supposing a uniform distribution, the probability of choosing such an integer is  $\frac{41}{1000}$ .

### Problem 14.

*Solution.* This is a weaker version of the [Borel-Cantelli lemma](#). Consider the Lebesgue probability measure on  $[0, 1]$ . And then consider the events  $\left[0, \frac{1}{n}\right]$  for  $n = 1, 2, \dots$ . It follows that  $\sum_{k=1}^{\infty} \mathbb{P} \left[ \left[0, \frac{1}{n}\right] \right] < \infty$ , but we know that the limit superior of the probability of this sequence of events is 0. Such a sequence of events cannot be described by the stronger version of the Borel-Cantelli lemma, however, the preconditions in this problem allow for this “counter example.”

First, we show the following lemma:

**Lemma 1.1.** For any indexed collection of sets  $C_1, C_2, \dots, C_n$ ,

$$\bigcup_{k=1}^n C_k = C_1 \cup \bigcup_{k=1}^{n-1} C_{k+1} \setminus C_k.$$

*Proof.* We proceed by way of induction. Suppose that  $n = 2$ , then  $C_1 \cup (C_2 \setminus C_1) = C_1 \cup (C_1^c \cap C_2) = (C_1 \cup C_1^c) \cap (C_2 \cup C_1) = C_1 \cup C_2$ . Assuming the claim true for all  $n < N$ , we now prove it for  $n = N$ :

$$\begin{aligned} C_1 \cup \bigcup_{k=1}^{n-1} (C_{k+1} \setminus C_k) &= \left( C_1 \cup \bigcup_{k=1}^{n-2} (C_{k+1} \setminus C_k) \right) \cup (C_n \setminus C_{n-1}), \\ &= \left( \bigcup_{k=1}^{n-1} C_k \right) \cup (C_{n-1}^c \cap C_n), && \text{(by ind. hyp.)} \\ &= \left( \bigcup_{k=1}^{n-2} C_k \right) \cup (C_{n-1} \cup (C_{n-1}^c \cap C_n)), \\ &= \left( \bigcup_{k=1}^{n-2} C_k \right) \cup (C_n \cup C_{n-1}), \\ &= \bigcup_{k=1}^n C_k. \end{aligned}$$

□

With this, we show the desired result:

$$\begin{aligned}
\mathbb{P} \left[ \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} C_k \right] &= \lim_{n \rightarrow \infty} \mathbb{P} \left[ \bigcup_{k=n}^{\infty} C_k \right], \\
&= \lim_{n \rightarrow \infty} \mathbb{P} \left[ C_n \cup \bigcup_{k=n}^{\infty} C_{k+1} \setminus C_k \right], \\
&\leq \lim_{n \rightarrow \infty} \mathbb{P}[C_n] + \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mathbb{P}[C_{k+1} \setminus C_k], \\
&= 0.
\end{aligned}$$

The second equality follows from Lemma 1.1 and the last equality follows from the given.

**Problem 15.**

*Solution.* If  $F$  is a continuous distribution of a random variable  $\xi$  on the probability space  $(\Omega, \mathcal{G}, \mathcal{P})$ , then it must be that  $F$  is injective on  $\text{img } \xi$ . Consider, now a random variable  $\eta = F(\xi)$ . We proceed to find the distribution of this  $\eta$ :

$$\begin{aligned}
\mathbb{P}[\{\omega \in \Omega : \eta(\xi(\omega)) \leq x\}] &= \mathbb{P}[\{\omega \in \Omega : \xi(\omega) \leq F^{-1}(x)\}], \\
&= F(F^{-1}(x)).
\end{aligned}$$

Note that unless  $F$  is bijective,  $F \circ F^{-1} \neq \text{Id}$ . Recall, that there exists a right-inverse if and only if  $F$  is surjective.

# Chapter 2

## Sequences of Independent Trials

Some brief notes before the problems:

- Koralov and Sinai use  $\times$  instead of  $\otimes$  for the product  $\sigma$ -algebra. Moving forward, the reader needs to be cautious about what the operands are. For if they are  $\sigma$ -algebras, we understand the result of the expression to be the aforementioned product.

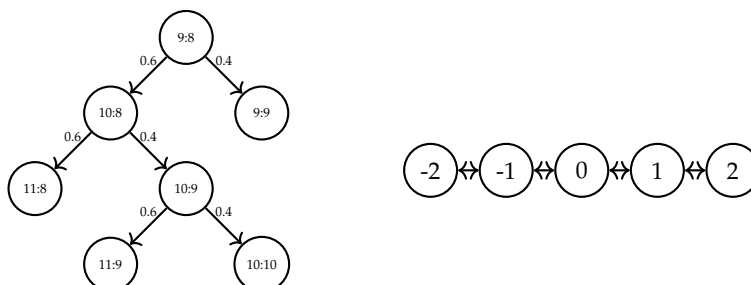
### Problem 1.

*Solution.* Directly apply the definition of the binomial distribution. Let  $X = \{0, 1\}$ . Define the random variable  $\chi_i(\omega)$  for  $1 \leq i \leq 5$  to be 1 if  $\omega = 1$  and zero otherwise. Then, let  $\nu_5 = \sum_{i=1}^5 \chi_i$ .  $\nu_5$  is a binomial random variable. Thus,

$$\mathbb{P}[\nu_5 = 3] = \binom{5}{3} \frac{1}{2^5} = \frac{5}{16}.$$

### Problem 2.

*Solution.* We present two alternative ways to solve this problem one involving chains and the other geometric series. We first present the solution involving chains. We draw the state space below, where each node denotes a distinct score and the edges between states represent the transition probabilities



It remains to find the probability that Andrew wins from either 9:9 or 10:10 which is the same since both states require Andrew to win two consecutive points. This, then induces a new chain (see above). Denote by  $a, b, c$  the probability that Andrew wins from being one point down, drawn, and one point up. Then,

$$a = 0.6b,$$



$$\begin{aligned} b &= 0.4a + 0.6c, \\ c &= 0.4b + 0.6b. \end{aligned}$$

So,  $c = 0.36/0.52$ . Thus, the probability that Andrew wins is then the aggregate of all states in the chain multiplied by their transition probabilities:

$$\mathbb{P}[\text{Andrew wins}] = 0.4 \left( \frac{0.36}{0.52} \right) + 0.6 \left( 0.6 + 0.4 \left( 0.4 \left( \frac{0.36}{0.52} \right) + 0.6 \right) \right) = 0.847.$$

### Problem 3.

*Solution.* We directly apply the de Moivre-Laplace Theorem. Let  $X_i(\omega_i)$  be a random variable that is 1 if  $\omega_i = \text{heads}$  and 0 otherwise. Then, let  $\nu_n = \sum_{i=1}^n X_i$ . Then, for  $n = 1000$ ,  $\mathbb{E}[\nu_n] = 500$  and  $\text{Var } \nu_n = 250$ .  $k = 600$  as the actual number of heads that we saw during all of the coin flips. Therefore,  $z = (600 - 500)/\sqrt{250} = 20/\sqrt{10}$ . Therefore, by the de Moivre-Laplace theorem:

$$\begin{aligned} \mathbb{P} \left[ \frac{\nu_n - \mathbb{E}[\nu_n]}{\sqrt{\text{Var}(\nu_n)}} \geq z \right] &\geq \int_{20/\sqrt{10}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} z^2 \right) dz, \\ &\leq 10^{-5}. \end{aligned}$$

So, it is highly unlikely that the coin is fair.

### Problem 4.

*Solution.* We show that  $\mathbb{P} \left[ \left| \frac{\nu^n}{n} - p \right| > \varepsilon_n \right] \rightarrow 0$  as  $n \rightarrow \infty$ :

$$\begin{aligned} \mathbb{P} \left[ \left| \frac{\nu^n}{n} - p \right| > \varepsilon_n \right] &= \mathbb{P}[|\nu^n - np| > n\varepsilon_n], \\ &\leq \frac{\text{Var } \nu^n}{n^2 \varepsilon_n^2}, \\ &= \frac{np(1-p)}{n(n\varepsilon_n)^2}, \\ &= \frac{p(1-p)}{(\sqrt{n}\varepsilon_n)^2} \rightarrow 0, \end{aligned}$$

as  $n \rightarrow \infty$  since  $\sqrt{n}\varepsilon_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

*Remark:* Even though this problem is fairly simple, it demonstrates something profound about the law of large numbers — its convergence rate. Specifically, the probability that  $\nu^n$  lies in a neighborhood of  $O(\sqrt{n})$  tends to 1, as  $n \rightarrow \infty$ . Moreover, the probability that  $\nu^n = k$  decays with rate  $O(1/\sqrt{n})$  if  $k \in (np - O(\sqrt{n}), np + O(\sqrt{n}))$ . This is easy to see since we can approximate the distribution within this neighborhood with a crude uniform distribution. Note that this is the exact setup of the De Moivre-Laplace Theorem.

### Problem 5.

*Solution.* Let  $\nu^n$  be the number of sixes that appear from  $n$  tosses of a die. Then, it follows that  $\mathbb{E}[\nu^{12000}] = 2000$  and  $\text{Var}[\nu^{12000}] = 12000 \cdot \frac{5}{36} \approx 1666$ . Using the de Moivre-Laplace Theorem, we approximate  $\frac{\nu^n - \mathbb{E}[\nu^n]}{\sqrt{\text{Var}[\nu^n]}}$  as a standard normal distribution. Therefore,

$$\mathbb{P}[1900 \leq \nu^n \leq 2100] = \Phi(2.45) - \Phi(-2.45),$$

where  $\Phi(\cdot)$  is the distribution function of the normal distribution.

**Problem 6.**

*Solution.*

$$\begin{aligned}
\mathbb{P}[\eta_n \leq t] &= 1 - \mathbb{P}[\eta_n > t], \\
&= 1 - \mathbb{P}[\omega_i \geq t]^n, \quad (\text{by i.i.d.}) \\
&= 1 - (1 - t)^n.
\end{aligned}$$

Next, we evaluate the limit as  $n \rightarrow \infty$  of  $n\eta_n$ :

$$\begin{aligned}
\lim_{n \rightarrow \infty} \mathbb{P}[n\eta_n \leq t] &= \lim_{n \rightarrow \infty} 1 - \mathbb{P}[n\eta_n > t], \\
&= \lim_{n \rightarrow \infty} 1 - \mathbb{P}\left[\eta_n > \frac{t}{n}\right], \\
&= \lim_{n \rightarrow \infty} 1 - \left(1 - \frac{t}{n}\right)^n, \\
&= 1 - e^{-t}.
\end{aligned}$$

**Problem 7.**

*Solution.*

**Problem 8.**

*Solution.* Let  $\Omega$  be the space of all binary sequences of length  $n$  and also let  $C_r$  be the set of all binary sequences of length  $n$  with  $r$  consecutive 1s. Then,

$$C_r = \bigsqcup_{i=1}^{n-r+1} \{\sigma \in \Omega : \sigma \text{ has } r \text{ consecutive 1s and is prefixed with } i-1 \text{ zeros}\}.$$

Clearly, these sets are disjoint and contain any sequence that has  $r$  consecutive 1s. Therefore,

$$\begin{aligned}
\mathbb{P}[C_r] &= \sum_{i=1}^{n-r+1} \mathbb{P}[\{\sigma \in \Omega : \sigma \text{ has } r \text{ consecutive 1s and is prefixed with } i-1 \text{ zeros}\}], \\
&= \sum_{i=1}^{n-r+1} \frac{1}{2^{r+i-1}}.
\end{aligned}$$

**Problem 9.**

*Solution.* Let  $\Omega$  be the space of all sequences  $(\omega_1, \dots, \omega_n)$  of length  $n$  such that  $\omega_i \in \{1.03, 0.97\}$  chosen uniformly. Suppose that the stock price is  $A$ , then the stock price after a sequence of percent changes

$(\omega_1, \dots, \omega_n)$  is  $A \cdot \prod_{i=1}^n \omega_i$ . In order to apply de Moivre-Laplace, we need a sum of Bernoulli random variables.

Thus, we naturally proceed to apply  $\exp(\log(\cdot))$ :

$$\begin{aligned}
\exp\left(\log\left(A \prod_{i=1}^n \omega_i\right)\right) &= \exp\left(\log A + \sum_{i=1}^n \log(\omega_i)\right), \\
&= A \exp\left(\sum_{i=1}^n \log(\omega_i)\right),
\end{aligned}$$

where in the last equality the term in the exponent is normally-distributed by the de Moivre-Laplace Theorem. We now derive this distribution. Let  $X \sim N(\mu, \sigma^2)$  and  $Z \sim \exp X$ :

$$\begin{aligned}\mathbb{P}[Z \leq z] &= \mathbb{P}[\exp X \leq z], \\ &= \mathbb{P}[X \leq \log(z)], \\ &= \Phi_{\mu, \sigma^2}(\log(z)),\end{aligned}$$

where  $\Phi_{\mu, \sigma^2}(\cdot)$  is the distribution function of a normally-distributed random variable with mean  $\mu$  and variance  $\sigma^2$ . Then, it follows that the density function is  $\frac{1}{z}p(\mu, \sigma^2; \log(z))$ , where  $p(\mu, \sigma; \cdot)$  is the density of a normally-distributed random variable.

It remains to derive the distribution of  $X$  using the de Moivre-Laplace Theorem. First, we redefine the sample space. Let  $\Omega'$  be the space of all sequences  $(\omega_1, \dots, \omega_n)$  of length  $n$  such that  $\omega_i \in \{\log(1.03), \log(0.97)\}$  chosen uniformly. Then,

$$\mathbb{E} \left[ \sum_{i=1}^n \omega_i \right] = \sum_{i=1}^n \mathbb{E} [\omega_i] = \frac{n}{2} (\log(0.9991)).$$

And

$$\begin{aligned}\text{Var} \left[ \sum_{i=1}^n \omega_i \right] &= \mathbb{E} \left[ \left( \sum_{i=1}^n \omega_i \right)^2 \right] - \mathbb{E} \left[ \sum_{i=1}^n \omega_i \right]^2, \\ &= \sum_{i=1}^n \mathbb{E} [\omega_i^2] + \sum_{1 \leq i \neq j \leq n} \mathbb{E} [\omega_i] \mathbb{E} [\omega_j] - \mathbb{E} \left[ \sum_{i=1}^n \omega_i \right]^2, \\ &= \frac{n}{2} (\log^2(1.03) + \log^2(0.97)) + \frac{n^2 - n}{4} \log^2(0.9991) - \frac{n^2}{4} \log^2(0.9991).\end{aligned}$$

Evaluating the mean and variance for  $n = 250$ , we have that  $X \sim N(-0.112551, 0.225135)$ . Applying the distribution function of the log-normal distribution with these parameters, we have  $\mathbb{P}[Z \geq 1] = 0.4111$ .

# Lebesgue Integral and Mathematical Expectation

## Problem 1.

*Solution.* It suffices to show that

$$\{\omega \in \Omega : \lim_{n \rightarrow \infty} f_n(\omega) = f(\omega)\} = \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} \left\{ \omega \in \Omega : |f_k(\omega) - f(\omega)| < \frac{1}{i} \right\}. \quad (3.1)$$

This is because

$$\left\{ \omega \in \Omega : |f_k(\omega) - f(\omega)| < \frac{1}{i} \right\} = (f_k - f)^{-1} \left( -\frac{1}{i}, \frac{1}{i} \right).$$

Since both  $f_k$  and  $f$  are  $\mathcal{F}$ -measurable functions, then  $f_k - f$  is also a measurable function. Moreover, since  $(-1/i, 1/i)$  is clearly a Borel set, then  $(f_k - f)^{-1}(-1/i, 1/i) \in \mathcal{F}$ . It follows from the axioms of a  $\sigma$ -algebra that the right hand-side of Eq. 3.1 is also a set in  $\mathcal{F}$ .

Now, we show Eq. 3.1:

$$\begin{aligned} x \in \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} \left\{ \omega \in \Omega : |f_k(\omega) - f(\omega)| < \frac{1}{i} \right\} &\Rightarrow x \in \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} \left\{ \omega \in \Omega : |f_k(\omega) - f(\omega)| < \frac{1}{i} \right\}, \quad (\forall i \in \mathbb{N}) \\ &\Rightarrow \exists j, \forall k > j, x \in \left\{ \omega \in \Omega : |f_k(\omega) - f(\omega)| < \frac{1}{i} \right\}, \quad (\forall i \in \mathbb{N}) \\ &\Rightarrow \forall i \in \mathbb{N}, \exists j, \forall k > j, |f_k(x) - f(x)| < \frac{1}{i}. \end{aligned}$$

It follows that  $f_1, f_2, \dots$  to  $f$  at  $x$ . The other direction of the inclusion is trivial and its proof is omitted as it results from the same logic above.

## Problem 2.

*Solution.*

$$\begin{aligned} \mathbb{E}[\xi] &\geq \sum_{k=1}^{\infty} k \cdot \mathbb{P}[k \leq \xi < k+1], \\ &= \mathbb{P}[1 \leq \xi < 2] + 2\mathbb{P}[2 \leq \xi < 3] + 3\mathbb{P}[3 \leq \xi < 4] + \dots, \\ &= \left( \sum_{k=1}^{\infty} \mathbb{P}[k \leq \xi < k+1] \right) + \mathbb{P}[2 \leq \xi < 3] + 2\mathbb{P}[3 \leq \xi < 4] + \dots, \end{aligned}$$

$$\begin{aligned}
&= 1 + \left( \sum_{k=2}^{\infty} \mathbb{P}[k \leq \xi < k+1] \right) + \dots, \\
&= 1 + \frac{1}{2} + \frac{1}{3} + \dots, \\
&\geq \infty.
\end{aligned}$$

### Problem 3.

*Solution.* Let  $\xi_n$  be a normally-distributed random variable parameterized by mean  $n$  and variance 1. Then, it is clear that  $\xi_n$  converges pointwise to the random variable  $\xi(\omega) = 0$ , however,

$$\lim_{n \rightarrow \infty} \mathbb{E}[\xi_n] = \lim_{n \rightarrow \infty} n = \infty.$$

### Problem 4.

*Solution.* Let  $\xi$  be a random variable such that  $\mathbb{P}[\xi = A] = \mathbb{P}[\xi = B] = \frac{1}{2}$ . Then it follows that

$$\begin{aligned}
\text{Var } \xi &= \mathbb{E}[\xi^2] - \mathbb{E}[\xi]^2, \\
&= \frac{1}{2}A^2 + \frac{1}{2}B^2 - \left( \frac{1}{2}A + \frac{1}{2}B \right)^2, \\
&= \frac{1}{2}A^2 + \frac{1}{2}B^2 - \frac{1}{4}(A+B)^2, \\
&= \frac{1}{4}A^2 - \frac{1}{2}AB + \frac{1}{4}B^2, \\
&= \left( \frac{B-A}{2} \right)^2.
\end{aligned}$$

Interestingly, this is the maximum variance a bounded random variable can achieve (see [Popoviciu's Inequality on Variances](#)).

### Problem 5.

*Solution.* Let  $x$  be an arbitrary irrational number and also let  $\varepsilon > 0$ . Then, choose  $N \in \mathbb{N}$  such that  $\log_2(1/\varepsilon) + 1 < N$ . Then, let

$$\delta = \min \{ |x - x_k| \}_{k=1}^N.$$

So, for any  $x'$  such that  $|x - x'| < \delta$ , it must be that

$$\begin{aligned}
|F_\xi(x) - F_\xi(x')| &= \left| \sum_{x_j \text{ s.t. } x_j < x} \frac{1}{2^j} - \sum_{x_j \text{ s.t. } x_j < x'} \frac{1}{2^j} \right|, \\
&= \sum_{x_j \text{ s.t. } x' \leq x_j < x} \frac{1}{2^j}.
\end{aligned}$$

This equality comes from assuming, without the loss of generality,  $x' < x$ . So, then

$$\begin{aligned}
&\leq \sum_{j=N}^{\infty} \frac{1}{2^j}, \\
&= \frac{1}{2^{N-1}}, \\
&< \varepsilon.
\end{aligned}$$

(by def.)

Thus,  $F_x i(\omega)$  is continuous for irrational  $\omega$ .

**Problem 6.**

*Solution.* We first find the density of  $\eta$ . Consider first the distribution of  $\eta$ :

$$\begin{aligned}\mathbb{P}[\eta \leq x] &= \mathbb{P}\left[\frac{1}{\xi} \leq x\right], \\ &= 1 - \mathbb{P}\left[\xi < \frac{1}{x}\right], \\ &= 1 - F_{\xi}\left(\frac{1}{x}\right).\end{aligned}$$

Therefore, the density function of  $\eta$ ,  $p_{\eta} = \left(1 - F_{\xi}\left(\frac{1}{x}\right)\right)' = \frac{1}{x^2}p_{\xi}\left(\frac{1}{x}\right)$ . Now, we apply the definition of the expected value of  $\xi$ :

**Problem 7.**

*Solution.* Here, we are essentially showing that convergence in  $L^1$  implies almost surely pointwise convergence.

**Problem 8.**

*Solution.* We first show that convergence almost surely implies convergence in measure.

**Problem 9.**

*Solution.* Apply integration by-parts:

$$\int_{-\infty}^{\infty} F(x+10) - F(x) dx = x(F(x+10) - F(x)) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} x(p(x+10) - p(x)) dx,$$

where  $p(x)$  is the density function corresponding to  $F(\cdot)$ . First, we evaluate the term on the left, since  $\lim_{x \rightarrow \infty} F(x) = 1$  and  $\lim_{x \rightarrow -\infty} F(x) = 0$ , it must be that  $x(F(x+10) - F(x)) \Big|_{-\infty}^{\infty} = 0$ . Therefore, it suffices to evaluate the term on the right,

$$\begin{aligned}- \int_{-\infty}^{\infty} x(p(x+10) - p(x)) dx &= - \int_{-\infty}^{\infty} xp(x+10) - xp(x) dx, \\ &= - \int_{-\infty}^{\infty} (x+10)p(x+10) - xp(x) dx - \int_{-\infty}^{\infty} 10p(x+10) dx, \\ &= - \int_{-\infty}^{\infty} up(u) du - \int_{-\infty}^{\infty} xp(x) + 10, \\ &= 10.\end{aligned}$$

In general, it follows that for any  $u \in \mathbb{R}$ ,  $\int F(x+u) - F(x) dx = u$  through the same calculation that we just performed.

**Problem 10.**

*Solution.*

**Problem 11.**

*Solution.*

**Problem 12.**

*Solution.*

**Problem 13.**

*Solution.*

# Conditional Probabilities and Independence

## Problem 1.

*Solution.* We directly apply the definition of conditional probability:

$$\begin{aligned}
 \mathbb{P}[\omega_1 = 1 | \omega_1 + \dots + \omega_n = m] &= \frac{\mathbb{P}[\omega_1 = 1 \cap \omega_1 + \dots + \omega_n = m]}{\mathbb{P}[\omega_1 + \dots + \omega_n = m]}, \\
 &= \frac{\binom{n-1}{m-1} p^m (1-p)^{n-m}}{\binom{n}{m} p^m (1-p)^{n-m}}, \\
 &= \frac{\binom{n-1}{m-1}}{\binom{n}{m}}, \\
 &= \frac{m}{n}.
 \end{aligned}$$

## Problem 2.

*Solution.* We are looking for a distribution that is *memoryless*, both the exponential distribution and the Poisson distribution satisfy this property. Let  $\xi$  be a exponentially-distributed random variable, parameterized by  $\lambda$ . Then,  $\mathbb{P}[\xi > x] = e^{-\lambda x}$  and

$$\begin{aligned}
 \mathbb{P}[\xi > x + y | \xi > x] &= \frac{\mathbb{P}[\xi > x + y]}{\mathbb{P}[\xi > x]}, \\
 &= \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}}, \\
 &= e^{-\lambda y}.
 \end{aligned}$$

## Problem 3.

*Solution.* Let  $H^1, H^2$  denote the event that heads appears on the first and second toss, respectively. Then,

$$\mathbb{P}[H^2 | H^1] = \frac{\mathbb{P}[H^1 \cap H^2]}{\mathbb{P}[H^1]},$$

$$\begin{aligned}
&= \frac{1/2(1/2^2) + 1/2(0.6^2)}{1/2(1/2) + 1/2(0.6)}, \\
&= 0.55.
\end{aligned}$$

**Problem 4.**

*Solution.* Suppose that  $\xi, \eta$  are two random variables such that both of their images is  $\{a, b\}$ . We show then that if  $\mathbb{E}[\xi\eta] = \mathbb{E}[\xi]\mathbb{E}[\eta]$ , then  $\eta$  and  $\xi$  are independent. Let  $p, q$  be the probability that  $\xi = a, \eta = a$ , respectively. Then,

$$\begin{aligned}
\mathbb{E}[\xi]\mathbb{E}[\eta] &= (pa + (1-p)b)(qa + (1-q)b), \\
&= a^2pq + ab(p)(1-q) + ba(q)(1-p) + b^2(1-p)(1-q), \\
\mathbb{E}[\xi\eta] &= a\mathbb{P}[\eta = a, \xi = a] + ab\mathbb{P}[\eta = a, \xi = b] + ba\mathbb{P}[\eta = b, \xi = a] + b^2\mathbb{P}[\xi = b, \eta = b],
\end{aligned}$$

This implies that

$$\begin{aligned}
\mathbb{P}[\xi = a, \eta = a] &= pq = \mathbb{P}[\xi = a]\mathbb{P}[\eta = a], \\
\mathbb{P}[\xi = b, \eta = b] &= (1-q)(1-p) = \mathbb{P}[\xi = b]\mathbb{P}[\eta = b].
\end{aligned}$$

By Lemma 4.3, then events  $\xi = a, \eta = b$  and  $\eta = a, \xi = b$  are also independent.

**Problem 5.**

*Solution.*

**Problem 6.**

*Solution.*

**Problem 7.**

*Solution.*