Solutions

Chapter 2 Foundations of Probability

2.1 (Composing random elements) Show that if f is $\mathcal{F}/\mathcal{G}$-measurable and g is $\mathcal{G}/\mathcal{H}$-measurable for sigma algebras $\mathcal{F}$,$\mathcal{G}$ and $\mathcal{H}$ over appropriate spaces, then their composition, $g \circ f$ (defined the usual way: $(g \circ f)(\omega) = g(f(\omega)), \omega \in \Omega$), is $\mathcal{F}/\mathcal{H}$-measurable.

Proof. Since g is $\mathcal{G}/\mathcal{H}$-measurable, therefore $\forall C \in \mathcal{H}$, $\exists B = g^{-1}(C) \in \mathcal{G}$. Similarly, since f is $\mathcal{F}/\mathcal{G}$-measurable, $\forall B \in \mathcal{G}$, $\exists A = f^{-1}(B) \in \mathcal{F}$. Thus $\forall C \in \mathcal{H}$, $\exists A = f^{-1}(g^{-1}(C)) = (g \circ f)^{-1}(C) \in \mathcal{F}$ and the proof is complete.

□

2.2 Let $X_1, \dots, X_n$ be random variables on $(\Omega, \mathcal{F})$. Prove that $X = (X_1, \dots, X_n)$ is a random vector.

Proof. Since X_i is a random variable ($\forall i = 1, 2, ..., n$), it holds that X_i is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable, which means that $\forall B \in \mathcal{B}(\mathbb{R})$, $X_i^{-1}(B) \in \mathcal{F}$. We first prove that X is $\mathcal{F}/(\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \cdots \mathcal{B}(\mathbb{R}))$-measurable (totally n $\mathcal{B}(\mathbb{R})$s). $\forall A = A_1 \times A_2 \times \cdots \times A_n \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \cdots \mathcal{B}(\mathbb{R})$, $X^{-1}(A) = X_1^{-1}(A_1) \cap X_2^{-1}(A_2) \cap \cdots \cap X_n^{-1}(A_n) \in \mathcal{F}$, which holds since $X_i^{-1}(A_i) \in \mathcal{F}, \forall i = 1, 2, ..., n$ and $\mathcal{F}$ is a σ-algebra. Thus we conclude that X is $\mathcal{F}/(\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \cdots \mathcal{B}(\mathbb{R}))$-measurable.

By definition $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \cdots \mathcal{B}(\mathbb{R}))$ (totally n $\mathcal{B}(\mathbb{R})$s). And according to the property in 2.5(b), we can get that X is $\mathcal{F}/\mathcal{B}(\mathbb{R}^n)$-measurable, thus it is a random vector.

□

2.3 (Random variable induced σ-algebra) Let $\mathcal{U}$ be an arbitrary set and $(\mathcal{V}, \Sigma)$ a measurable space and $X : \mathcal{U} \to \mathcal{V}$ an arbitrary function. Show that $\Sigma_X = \{X^{-1}(A) : A \in \Sigma\}$ is a σ-algebra over $\mathcal{U}$.

Proof.

(i) We need to show that Σ_X is closed under countable union. Let $U_i = X^{-1}(A_i), A_i \in \Sigma, i \in \mathbb{N}$. It follows that $\bigcup\limits_{i=1}^{\infty} U_i = \bigcup\limits_{i=1}^{\infty} X^{-1}(A_i) = X^{-1}(\bigcup\limits_{i=1}^{\infty} A_i)$. Since $\bigcup\limits_{i=1}^{\infty} A_i \in \Sigma$, $\bigcup\limits_{i=1}^{\infty} U_i \in \Sigma_X$.

(ii) We need to show that Σ_X is closed under set subtraction $-$. $\forall U_1, U_2 \in \Sigma_X$,$U_1 - U_2 = X^{-1}(A_1) - X^{-1}(A_2) = X^{-1}(A_1 - A_2)$. Since $A_1 - A_2 \in \Sigma$, $U_1 - U_2 \in \Sigma_X$.

(iii) We need to show that Σ_X is closed to $\mathcal{U}$ itself. Since $\mathcal{U} = X^{-1}(\mathcal{V})$ and $\mathcal{V} \in \Sigma$, it follows that $\mathcal{U} \in \Sigma_X$.

□

2.4 Let $(\Omega, \mathcal{F})$ be a measurable space and $A \subseteq \Omega$ and $\mathcal{F}_{|A} = \{A \cap B : B \in \mathcal{F}\}$.

Proof. (a)

(i) We need to show that $\mathcal{F}|_A$ is closed under countable union. Let $X_1 = A \cap B_1, X_2 = A \cap B_2, ...$ and $X' = \bigcup\limits_{i=1}^{\infty} X_i$ and $B' = \bigcup\limits_{i=1}^{\infty} B_i$ where $B_1, B_2, ... \in \mathcal{F}$. Since $\mathcal{F}$ is sigma algebra, $B' \in \mathcal{F}$. Furthermore, since $X' = \bigcup_{i=1}^{\infty} X_i = \bigcup_{i=1}^{\infty} A \bigcap B_i = A \bigcap \left(\bigcup\limits_{i=1}^{\infty} B_i\right) = A \bigcap B'$, we can see that $X' \in \mathcal{F}|_A$.

(ii) We need to show that $\mathcal{F}|_A$ is closed under set subtraction $-$. $\forall X_1, X_2 \in \mathcal{F}|_A$, $X_1 - X_2 = (A \bigcap B_1) - (A \bigcap B_2) = A \bigcap (B_1 - B_2)$. Since $B_1 - B_2 \in \mathcal{F}$, it follows that $X_1 - X_2 \in \mathcal{F}|_A$.

(iii) We need to show that $\mathcal{F}|_A$ is closed to A itself. Since $\varnothing \in \mathcal{F}$, we have $\varnothing = A \bigcap \varnothing \in \mathcal{F}|_A$ and $A = \varnothing^C \in \mathcal{F}|_A$.

(b) Let $P = \{A \bigcap B : B \in \mathcal{F}\}, Q = \{B : B \subset A, B \in \mathcal{F}\}$.

(i) We claim that $P \subset Q$. Let $X = A \bigcap B, B \in \mathcal{F}$. Since $A \in \mathcal{F}$, $X = A \bigcap B \in \mathcal{F}$. Furthermore, $X \in Q = \{B : B \subset A, B \in \mathcal{F}\}$.

(ii) We claim that $Q \subset P$. $\forall X \in Q$, we have $X \subset A$ and $X \in \mathcal{F}$, which means that $X = X \bigcap A$ and $X \in \mathcal{F}$. It follows that $X \in P$.

(iii) Take both (i)(ii) into consideration, we can see that $P = Q$.

□

2.5 Let $\mathcal{G} \subseteq 2^\Omega$ be a non-empty collection of sets and define $\sigma(\mathcal{G})$ as the smallest σ-algebra that contains $\mathcal{G}$. By 'smallest' we mean that $\mathcal{F} \in 2^\Omega$ is smaller than $\mathcal{F}' \in 2^\Omega$ if $\mathcal{F} \subset \mathcal{F}'$.

(a) Show that $\sigma(\mathcal{G})$ exists and contains exactly those sets A that are in every σ-algebra that contains $\mathcal{G}$.

(b) Suppose $(\Omega', \mathcal{F})$ is a measurable space and $X : \Omega' \to \Omega$ be $\mathcal{F}/\mathcal{G}$-measurable. Show that X is also $\mathcal{F}/\sigma(\mathcal{G})$-measurable. (We often use this result to simplify the job of checking whether a random variable satisfies some measurability property).

(c) Prove that if $A \in \mathcal{F}$ where $\mathcal{F}$ is a σ-algebra, then $\mathbb{I}\{A\}$ is $\mathcal{F}$-measurable.

Proof. (a) Let $\mathcal{K} = \{\mathcal{F} | \mathcal{F}$ is a σ-algebra and contains $\mathcal{G}\}$, It holds obviously that $\mathcal{K}$ is not an empty set since it contains $2^\mathcal{G}$.

Then $\bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$ contains exactly those sets that are in every σ-algebra that contains $\mathcal{G}$. Given its existence, we only need to prove that $\bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$ is the smallest σ-algebra that contains $\mathcal{G}$.

First we show $\bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$ is a σ-algebra. Since $\mathcal{F}$ is a σ-algebra and therefore $\Omega \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$, it follows that $\Omega \in \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$. Next, for any $A \in \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$, $A^c \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Since they are all σ-algebras, $A^c \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Hence $A^c \in \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$. Finally, for any $\{A_i\}_i \subset \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$, $\{A_i\}_i \subset \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Since they are all σ-algebras, $\bigcup_i A_i \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Hence $\bigcup_i A_i \in \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$.

Next we want to prove $\bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$ is the smallest σ-algebra that contains $\mathcal{G}$. It is quite obvious that $\bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F} \subseteq \mathcal{F}'$ for all $\mathcal{F}' \in \mathcal{K}$.

Above all, we have $\sigma(\mathcal{G}) = \bigcap_{\mathcal{F}\in\mathcal{K}} \mathcal{F}$.

(b) Define $\mathcal{H} = \{A : X^{-1}(A) \in \mathcal{F}\}$. To show X is $\mathcal{F}/\sigma(\mathcal{G})$-measurable, it is sufficient to prove $\sigma(\mathcal{G}) \subseteq \mathcal{H}$.

First we prove that $\mathcal{H}$ is a σ-algebra. It holds that $\Omega \in \mathcal{H}$ since $X^{-1}(\Omega) = \Omega' \in \mathcal{F}$. For any $A \in \mathcal{H}$, we have $X^{-1}(A) \in \mathcal{F}$, thus $X^{-1}(A^c) = X^{-1}(A)^c \in \mathcal{F}$, which holds since $\mathcal{F}$ is a σ-algebra. Thus $A^c \in \mathcal{H}$. For any $A_i \in \mathcal{F}$, $i = 1, 2, ...$, $X^{-1}(A_i) \in \mathcal{F}$, $X^{-1}(\cup_i A_i) = \cup_i X^{-1}(A_i) \in \mathcal{F}$. We can then conclude $\cup_i A_i \in \mathcal{H}$ and $\mathcal{H}$ is a σ-algebra.

Also, since X is $\mathcal{F}/\mathcal{G}$-measurable, we have $\mathcal{G} \subseteq \mathcal{H}$. Thus $\mathcal{H}$ is σ-algebra that contains $\mathcal{G}$. By applying the result of (a), we have $\sigma(\mathcal{G}) \subseteq \mathcal{H}$, which completes the proof.

(c) The idea is to show $\forall B \in \mathfrak{B}(\mathbb{R})$, $\mathbb{I}\{A\}^{-1}(B) \in \mathcal{F}$.

If $\{0, 1\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = \Omega \in \mathcal{F}$. If $\{0\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = A^c \in \mathcal{F}$. If $\{1\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = A \in \mathcal{F}$. If $\{0, 1\} \cap B = \emptyset$, $\mathbb{I}\{A\}^{-1}(B) = \emptyset \in \mathcal{F}$.

□

2.6 (Knowledge and σ-algebras: a pathological example) In the context of Lemma 2.5, show an example where $Y = X$ and yet Y is not $\sigma(X)$ measurable.

Hint As suggested after the lemma, this can be arranged by choosing $\Omega = \mathcal{Y} = \mathcal{X} = \mathbb{R}, X(\omega) = Y(\omega) = \omega, \mathcal{F} = \mathcal{H} = \mathfrak{B}(\mathbb{R})$ and $\mathcal{G} = \{\emptyset, \mathbb{R}\}$ to be the trivial σ-algebra.

Proof. As the hint suggests, Let $\Omega = \mathcal{Y} = \mathcal{X} = \mathbb{R}, X(\omega) = Y(\omega) = \omega, \mathcal{F} = \mathcal{H} = \mathfrak{B}(\mathbb{R})$. In this case, $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{G}\} = \{\emptyset, \mathbb{R}\}$, we can find that $Y^{-1}((0,1)) = (0,1) \notin \sigma(X)$, thus Y is not $\sigma(X)$-measurable. □

2.7 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $B \in \mathcal{F}$ be such that $\mathbb{P}(B) > 0$. Prove that $A \mapsto \mathbb{P}(A|B)$ is a probability measure over $(\Omega, \mathcal{F})$.

Proof. First we have $\mathbb{P}(\Omega \mid B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$. Then, for any $A \in \mathcal{F}$, $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geq 0$. Next, for any $A \in \mathcal{F}$, $\mathbb{P}(A^c \mid B) = \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((\Omega - A) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B) - \mathbb{P}(A \cap B)}{\mathbb{P}(B)} = 1 - \mathbb{P}(A \mid B)$. Finally, for all countable collections of disjoint sets $\{A_i\}_i$ with $A_i \in \mathcal{F}$ for all i, we have $\mathbb{P}\left(\bigcup_i A_i \mid B\right) = \frac{\mathbb{P}((\bigcup_i A_i) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\bigcup_i (A_i \cap B))}{\mathbb{P}(B)} = \sum_i \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_i \mathbb{P}(A_i \mid B)$. □

2.8 (Bayes law) Verify (2.2).

Proof. With the definition of conditional probability, we have $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$. □

2.9 Consider the standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$ generated by two standard, unbiased, six-sided dice that are thrown independently of each other. Thus, $\Omega = \{1, ..., 6\}^2$, $\mathcal{F} = 2^\Omega$ and $\mathbb{P}(A) = |A|/6^2$ for any $A \in \mathcal{F}$ so that $X_i(\omega) = \omega_i$ represents the outcome of throwing dice $i \in \{1, 2\}$.

(a) Show that the events '$X_1 < 2$' and 'X_2 is even' are independent of each other.

(b) More generally, show that for any two events, $A \in \sigma(X_1)$ and $B \in \sigma(X_2)$, are independent of each other.

Proof. (a) The event $\{X_1 < 2\} = \{1\} \times \{1,2,3,4,5,6\}$, $\{X_2 \text{ is even }\} = \{1,2,3,4,5,6\} \times \{2,4,6\}$, $\{X_1 < 2, X_2 \text{ is even }\} = \{(1,2),(1,4),(1,6)\}$.
Thus $\mathbb{P}(X_1 < 2) = \frac{6}{36} = \frac{1}{6}$, $\mathbb{P}(X_2 \text{ is even }) = \frac{18}{36} = \frac{1}{2}$, $\mathbb{P}(X_1 < 2, X_2 \text{ is even }) = \frac{3}{36} = \frac{1}{12}$, which satisfies $\mathbb{P}(X_1 < 2, X_2 \text{ is even }) = \mathbb{P}(X_1 < 2) \times \mathbb{P}(X_2 \text{ is even })$. These two events are independent of each other.

(b) $\sigma(X_1) = \left\{X_1^{-1}(A'), A' \subseteq [6]\right\} = \{A' \times [6] : A' \subseteq [6]\}, \sigma(X_2) = \left\{X_2^{-1}(B'), B' \subseteq [6]\right\} = \{[6] \times B' : B' \subseteq [6]\}$.
Thus $\forall A \in \sigma(X_1), B \in \sigma(X_2)$, $\mathbb{P}(A) = \frac{|A'| \times 6}{36} = \frac{|A'|}{6}$, $\mathbb{P}(B) = \frac{6 \times |B'|}{36} = \frac{|B'|}{6}$ and $\mathbb{P}(A \cap B) = \frac{|A'| \times |B'|}{36} = \mathbb{P}(A) \times \mathbb{P}(B)$. So A and B are independent of each other. □

2.10 (Serendipitous independence) The point of this exercise is to understand independence more deeply. Solve the following problems:

(a) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Show that $\emptyset$ and Ω (which are events) are independent of any other event. What is the intuitive meaning of this?

(b) Continuing the previous part, show that any event $A \in \mathcal{F}$ with $\mathbb{P}(A) \in \{0, 1\}$ is independent of any other event.

(c) What can we conclude about an event $A \in \mathcal{F}$ that is independent of its complement, $A^c = \Omega \setminus A$? Does your conclusion make intuitive sense?

(d) What can we conclude about an event $A \in \mathcal{F}$ that is independent of itself? Does your conclusion make intuitive sense?

(e) Consider the probability space generated by two independent flips of unbiased coins with the smallest possible σ-algebra. Enumerate all pairs of events A, B such that A and B are independent of each other.

(f) Consider the probability space generated by the independent rolls of two unbiased three-sided dice. Call the possible outcomes of the individual dice rolls 1, 2 and 3. Let X_i be the random variable that corresponds to the outcome of the ith dice roll ($i \in \{1, 2\}$). Show that the events $\{X_1 \le 2\}$ and $\{X_1 = X_2\}$ are independent of each other.

(g) The probability space of the previous example is an example when the probability measure is uniform on a finite outcome space (which happens to have a product structure). Now consider any n-element, finite outcome space with the uniform measure. Show that A and B are independent of each other if and only if the cardinalities $|A|$, $|B|$, $|A \cap B|$ satisfy $n|A \cap B| = |A| \cdot |B|$.

(h) Continuing with the previous problem, show that if n is prime, then no non-trivial events are independent (an event A is **trivial** if $\mathbb{P}(A) \in \{0, 1\}$).

(i) Construct an example showing that pairwise independence does not imply mutual independence.

(j) Is it true or not that A, B, C are mutually independent if and only if $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$? Prove your claim.

Proof. (a) Empty sets and complete sets are independent of any event:

$$P(A \cap \Omega) = P(A) = 1 \times P(A) = P(\Omega) \times P(A)$$
$$P(A \cap \emptyset) = P(\emptyset) = 0 = P(\emptyset) \times P(A)$$

(b) For any $B \in \Omega$ and $P(A) \in \{0, 1\}$:
when $P(A) = 1$,$P(A^c \cap B) \le P(A^c) = 1 - P(A) = 0$, we have $P(A \cap B) = P(A \cap B) + P(A^c \cap B) = P(B) = P(A)P(B)$; when $P(A) = 0$,we have $P(A \cap B) \le P(A) = 0 = P(A)P(B)$

(c) $P(A^c \cap A) = P(A)P(A^c)$, we have $0 = P(A)(1 - P(A)) \Rightarrow P(A) \in \{0, 1\}$

(d) $P(A \cap A) = P(A)P(A)$, we have $P(A) = \{0, 1\}$

(e) $\Omega = \{(1, 1), (1, 0), (0, 1), (0, 0)\}.A, B \subseteq \Omega$ denote the events.

First of all, if either A or B is trival, then A and B are independent of each other.

Then, we only need to enumerate $A, B \notin \Omega, \emptyset$ satisfied that $P(A \bigcap B) = P(A)P(B)$. Since $P(A \bigcap B) = \frac{|A \bigcap B|}{|\Omega|} = \frac{|A \bigcap B|}{4}$ and $P(A)P(B) = \frac{|A||B|}{16}$, we can conclude that $|A| = 2$, $|B| = 2$ and $|A \bigcap B| = 1$ is the only situation satisfying the condition.

Thus, besides trival A or B, all A, B satisfying $|A| = 2$, $|B| = 2$ and $|A \bigcap B| = 1$ are the solution.

(f) $P(X_1 \le 2) = 2/3$

$P(X_1 = X_2) = 3/9 = 1/3$

$P(X_1 \le 2, X_1 = X_2) = P(X_1 = X_2 = 1) + P(X_1 = X_2 = 2) = 1/9 + 1/9 = 2/9$

So, $P(X_1 \le 2, X_1 = X_2) = P(X_1 = X_2)P(X_1 \le 2)$

(g) Necessity :$\frac{|A \bigcap B|}{n} = P(A \bigcap B) = P(A)P(B) = \frac{|A|}{n}\frac{|B|}{n}$

$\Rightarrow |A \bigcap B| \times n = |A||B|$

Sufficiency :$|A \bigcap B| \times n = |A||B| \Rightarrow \frac{|A|}{n}\frac{|B|}{n} = \frac{|A \bigcap B|}{n}$

$\Rightarrow P(A \bigcap B) = P(A)P(B)$

(h) If A, B are two non-trival events independent to each other, $|A \bigcap B| \times n = |A||B| \Rightarrow n| \ (|A||B|) \Rightarrow n| \ (|A|)$ or $n| \ (|B|) \Rightarrow |A| = n$ or $|B| = n$, contradictory to non-trival assumption.

(i) Let $\Omega = \{1, 2, 3, 4\}$, $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$. A, B, C are pairwise independent but $P(A \bigcap B \bigcap C) = \frac{1}{4} \ne P(A)P(B)P(C) = \frac{1}{8}$.

(j) Consider rolling a dice and set $A = \{1,2,3\}$, $B = \{1,2,4\}$,$C = \{1,4,5,6\}$. Then $P(A \cap B \cap C) = \frac{1}{6} = (1/2) * (1/2) * (2/3) = P(A)P(B)P(C)$, however $P(A \cap B) = 1/3 \neq \frac{1}{2} * \frac{1}{2} = P(A)P(B)$. Thus $P(A \cap B \cap C) = P(A)P(B)P(C)$ does not mean mutuall independence.

$\square$

2.11 (Independence and random elements) Solve the following problems:

(a) Let X be a constant random element (that is, $X(\omega) = x$ for any $\omega \in \Omega$ over the outcome space over which X is defined). Show that X is independent of any other random variable.

(b) Show that the above continues to hold if X is almost surely constant (that is, $\mathbb{P}(X = x) = 1$ for an appropriate value x).

(c) Show that two events are independent if and only if their indicator random variables are independent (that is, A, B are independent if and only if $X(\omega) = \mathbb{1}\{\omega \in A\}$ and $Y(\omega) = \mathbb{1}\{\omega \in B\}$ are independent of each other).

(d) Generalise the result of the previous item to pairwise and mutual independence for collections of events and their indicator random variables.

Proof. (a) To prove X is independent of another random variable Y, we can equivalently show that $\sigma(X)$ and $\sigma(Y)$ are independent. And notice $\sigma(X) = \{\emptyset, \Omega\}$ for constant random element X. Therefore, for all $A \in \sigma(X)$ and $B \in \sigma(Y)$ it trivially holds that $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

(b) Given that $\mathbb{P}(X = x) = 1$, we can infer the generated sigma-algebra $\sigma(X) = \{\emptyset, \Omega, G_1, G_2, \cdots\}$, where $\mathbb{P}(G_1) = \mathbb{P}(G_2) = \cdots = 0$. Therefore, for any $A \in \sigma(X)$, we have $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$. By the result of 2.10(b), X is independent of any other random variable.

(c) Notice that $\sigma(X) = \{\emptyset, \Omega, A, A^c\}$, $\sigma(Y) = \{\emptyset, \Omega, B, B^c\}$.

(i) 'only if': Given that A, B are independent, we have $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Consequently $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B^c)$, which implies that A and B^c are also independent. Given that $\emptyset$ and Ω are trivially independent of any other event, we have A and B are independent for all $A \in \sigma(X)$ and $B \in \sigma(Y)$.

(ii) 'if': If X and Y are independent, $A \in \sigma(X)$ and $B \in \sigma(Y)$ are trivially independent.

(d) Notice that $\sigma(X_i) = \{\emptyset, \Omega, A_i, A_i^c\}$.

(i) Pairwise independence: The result can be generalized as we go through all pair of events.

(ii) Mutual independence: 'if' case is again trivial. For 'only if' case, suppose that $(A_i)_i$ are mutually independent. The mutual independence suggests that for any finite subset $K \subset \mathbb{N}$ we have$\mathbb{P}\left(\bigcap_{i \in K} A_i\right) = \prod_{i \in K} \mathbb{P}(A_i)$.
Similar to the previous part, for any disjoint finite sets J, K we have

$$\mathbb{P}\left(\bigcap_{i \in K} A_i \cap \bigcap_{i \in J} A_i^c\right) = \prod_{i \in K} \mathbb{P}(A_i) \prod_{i \in J} \mathbb{P}(A_i^c).$$

This leads to the conclusion that for any finite set $K \subset \mathbb{N}$ and $(V_i)_{i \in K}$ with $V_i \in \sigma(X_i) = \{\Omega, \emptyset, A_i, A_i^c\}$, we have

$$\mathbb{P}\left(\bigcap_{i \in K} V_i\right) = \prod_{i \in K} \mathbb{P}(V_i),$$

which implies that $(X_i)_i$ are mutually independent.

$\square$

2.12 Our goal in this exercise is to show that X is integrable if and only if $|X|$ is integrable. This is broken down into multiple steps. The first issue is to deal with the measurability of $|X|$. While a direct calculation can also show this, it may be worthwhile to follow a more general path:

(a) Any $f : \mathbb{R} \to \mathbb{R}$ continuous function is Borel measurable.

(b) Conclude that for any random variable X, $|X|$ is also a random variable.

(c) Prove that for any random variable X, X is integrable if and only if $|X|$ is integrable. (The statement makes sense since $|X|$ is a random variable whenever X is).

Proof. (a) Let $\mathcal{G} = \{(a,b) : a < b \text{ with } a, b \in \mathbb{R}\}$, then $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{G})$. According to Exercise 2.5(b), to show that f is $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable, we just need to show f is $\mathcal{B}(\mathbb{R})/\mathcal{G}$-measurable. Recall the definition of continuous function, $\forall x_0 \in \mathbb{R}, \varepsilon > 0, \exists \delta > 0$ such that when $x \in (x_0 - \delta, x_0 + \delta)$, there is $f(x) \in (f(x_0)-\varepsilon, f(x_0)+\varepsilon)$. Thus $\forall (a,b) \in \mathcal{G}, y_0 \in (a,b), y \in (y_0-\varepsilon, y_0+\varepsilon)$: $f^{-1}(y) \in (f^{-1}(y_0)-\delta, f^{-1}(y_0)+\delta)$, which means $f^{-1}(a,b) = \cup(a', b') \in \mathcal{B}(\mathbb{R})$. We have then shown that f is $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable.

(b) By definition, X is a random variable means that X is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable. According to Exercise 2.12 (a), $|X|$ is continous, thus it is $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable. Further apply the result of Exercise 2.1, let $f(X) = |X|$, then f is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable, thus $|X|$ is also a random variable.

(c) Define $X^+(\omega) = X(\omega)\mathbb{1}\{X(\omega) > 0\}$, $X^-(\omega) = -X(\omega)\mathbb{1}\{X(\omega) < 0\}$. If X is integrable, $\int_\Omega X d\mathbb{P} = \int_\Omega X^+ d\mathbb{P} - \int_\Omega X^- d\mathbb{P}$, which means X^+ and X^- are integrable, and by definition $\int_\Omega |X| d\mathbb{P} = \int_\Omega X^+ d\mathbb{P} + \int_\Omega X^- d\mathbb{P}$ is also integrable, vice versa.

□

2.13 (Infinite-valued integrals) Can we consistently extend the definition of integrals so that for non-negative random variables, the integral is always defined (it may be infinite)? Defend your view by either constructing an example (if you are arguing against) or by proving that your definition is consistent with the requirements we have for integrals.

Proof. We can extend the definition by letting at least one of $\int_\Omega X^+ d\mathbb{P}$ and $\int_\Omega X^- d\mathbb{P}$ be finite. □

2.14 Prove Proposition 2.6 (Let $(X_i)_i$ be a (possibly infinite) sequence of random variables on the same probability space and assume that $\mathbb{E}[X_i]$ exists for all i and furthermore that $X = \sum_i X_i$ and $\mathbb{E}[X]$ also exists. Then $\mathbb{E}[X] = \sum_i \mathbb{E}[X_i]$)

Proof. We only consider the infinite condition, as when the sum is finite, we can use the property of integration to draw the conclusion. We should add a condition first, which is

$$\sum_i E|X_i| < \infty \tag{1}$$

First, we need to prove

$$\mathbb{E}[\sum_i |X_i|] = \sum_i \mathbb{E}[|X_i|] \tag{2}$$

Consider $Y_n = \sum_{i=1}^n |X_i|$. As $\{Y_n\}_{n=1}^\infty$ is an ascent sequence and Y_n $\forall n$ is integrable. According to the monotone convergence theorem, we have

$$\mathbb{E}[\lim_{n\to\infty} Y_n] = \lim_{n\to\infty} \mathbb{E}[Y_n] \tag{3}$$

which is equivalent to (2).

Then we consider $Z_n = \sum_{i=1}^n X_i$, as $|Z_n| \le \sum_{i=1}^n |X_i| \le \sum_{i=1}^\infty |X_i|$. According to (1), $\sum_{i=1}^\infty |X_i|$ is integrable. Then, using dominated convergence theorem, we have

$$\mathbb{E}[\lim_{n\to\infty} Z_n] = \lim_{n\to\infty} \mathbb{E}[Z_n] \iff \mathbb{E}[X] = \sum_i \mathbb{E}[X_i] \tag{4}$$

□

2.15 Prove that if $c \in \mathbb{R}$ is a constant, then $\mathbb{E}[cX] = c\mathbb{E}[X]$ (as long as X is integrable).

Proof. (a) Firstly, assume X is a simple function and $X(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{I}\{\omega \in A_i\}$. Then, we have $cX(\omega) = \sum_{i=1}^{n} c\alpha_i \mathbb{I}\{\omega \in A_i\}$ and

$$\begin{aligned}\mathbb{E}[cX] &= \sum_{i=1}^{n} c\alpha_i \mathbb{P}(A_i) \\ &= c\sum_{i=1}^{n} \alpha_i \mathbb{P}(A_i) \\ &= c\mathbb{E}[X].\end{aligned}$$

(b) Assume $X \geq 0$ and $c \geq 0$. In this case,

$$\begin{aligned}\mathbb{E}[cX] &= \sup\{\mathbb{E}[h] : h \text{ is simple and } h \leq cX\} \\ &= \sup\{\mathbb{E}[ch] : h \text{ is simple and } h \leq X\} \\ &= \sup\{c\mathbb{E}[h] : h \text{ is simple and } h \leq X\} \\ &= c\sup\{\mathbb{E}[h] : h \text{ is simple and } h \leq X\} \\ &= c\mathbb{E}[X].\end{aligned}$$

(c) Assume X is arbitrary and $c \geq 0$. By factorizing $X = X^+ + X^-$ we can deduce that

$$\begin{aligned}\mathbb{E}[cX] &= \mathbb{E}[cX^+] - \mathbb{E}[cX^-] \\ &= c\mathbb{E}[X^+] - c\mathbb{E}[X^-] \\ &= c\mathbb{E}[X^+ - X^-] \\ &= c\mathbb{E}[X].\end{aligned}$$

(d) For $c < 0$ we may repeat the above procedures by only noticing that $(cX)^+ = -cX^-$ and $(cX)^- = -cX^+$.

□

2.16 Prove Proposition 2.7.

Hint Follow the 'inductive' definition of Lebesgue integrals, starting with simple functions, then non-negative functions and finally arbitrary independent random variables.

Proof. If X and Y are independent, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for any $A \in \sigma(X)$ and $B \in \sigma(Y)$.

(a) Assume X and Y are both simple. Furthermore, we can construct $X(\omega) = \sum_{i=1}^{m} \alpha_i \mathbb{I}\{\omega \in A_i\}$ and $Y(\omega) = \sum_{j=1}^{n} \beta_j \mathbb{I}\{\omega \in B_j\}$ where $(A_i)_i$ and $(B_j)_j$ are disjoint. In such a case, we have

$$\begin{aligned}\mathbb{E}[XY] &= \mathbb{E}[\sum_{i=1}^{m} \alpha_i \mathbb{I}\{A_i\} \sum_{j=1}^{n} \beta_j \mathbb{I}\{B_j\}] \\ &= \mathbb{E}[\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \mathbb{I}\{A_i\}\mathbb{I}\{B_j\}] \\ &= \mathbb{E}[\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \mathbb{I}\{A_i \cap B_j\}] \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \mathbb{P}\{A_i \cap B_j\} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \mathbb{P}\{A_i\}\mathbb{P}\{B_j\} \\ &= \sum_{i=1}^{m} \alpha_i \mathbb{P}\{A_i\} \sum_{j=1}^{n} \beta_j \mathbb{P}\{B_j\} \\ &= \mathbb{E}[X]\mathbb{E}[Y],\end{aligned}$$

where the fifth equation follows from $A_i \in \sigma(X)$ and $B_j \in \sigma(Y)$, provided that $(A_i)_i$ and $(B_j)_j$ are disjoint respectively.

(b) Assume X and Y are non-negative. Then, we can construct two monotone increasing sequences of simple functions $(g_i)_i$ and $(h_i)_i$ for X and Y respectively, where $\lim_{i\to\infty} g_i = X$ and $\lim_{i\to\infty} h_i = Y$ and $\mathbb{E}[g_i h_i] = \mathbb{E}[g_i]\mathbb{E}[h_i]$ following the same procedure in (a). Such sequences of simple functions can be constructed by making piecewise functions finer. Finally, by applying the monotone convergence theorem we have

$$\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}[\lim_{i\to\infty} g_i h_i] \\
&= \lim_{i\to\infty} \mathbb{E}[g_i h_i] \\
&= \lim_{i\to\infty} \mathbb{E}[g_i]\mathbb{E}[h_i] \\
&= \mathbb{E}[\lim_{i\to\infty} g_i]\mathbb{E}[\lim_{i\to\infty} h_i] \\
&= \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

(c) For arbitrary X and Y, we separate them into $X = X^+ - X^-$ and $Y = Y^+ - Y^-$, deducing that

$$\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}[(X^+ - X^-)(Y^+ - Y^-)] \\
&= \mathbb{E}[X^+Y^+] + \mathbb{E}[X^+Y^-] + \mathbb{E}[X^-Y^+] + \mathbb{E}[X^-Y^-] \\
&= \mathbb{E}[X^+]\mathbb{E}[Y^+] + \mathbb{E}[X^+]\mathbb{E}[Y^-] + \mathbb{E}[X^-]\mathbb{E}[Y^+] + \mathbb{E}[X^-]\mathbb{E}[Y^-] \\
&= \mathbb{E}[X^+ - X^-]\mathbb{E}[Y^+ - Y^-] \\
&= \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

□

2.17 Suppose that $\mathcal{G}_1 \subseteq \mathcal{G}_2$ and prove that $\mathbb{E}[X \mid \mathcal{G}_1] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}_1] \mid \mathcal{G}_2]$ almost surely.

Proof. The implication of this problem is to show that if X is $\mathcal{G}$-mearuable, then $\mathbb{E}[X \mid \mathcal{G}] = X$. Recall the definition of the conditional expectation.

Definition 1. *Assume $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-algebra of $\mathcal{F}$. $X : \Omega \to \mathbb{R}$ is a random variable. The conditional expectation of X given $\mathcal{G}$ is denoted by any random variable Y which satisfies the following 2 properties:*

- *Y is $\mathcal{G}$-measurable*
- *$\forall A \in \mathcal{G}$,*

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$$

Formally, we denoted Y by notation $\mathbb{E}[X|\mathcal{G}]$.

Since X is $\mathcal{G}$-measurable, property1 holds. And property2 holds trivially. Thus we can conclude the implication, and then the result of this problem holds trivially as $\mathbb{E}[X \mid \mathcal{G}_1]$ is $\mathcal{G}_1$-mearuable and thus $\mathcal{G}_2$-mearuable. □

2.18 Demonstrate using an example that in general, for dependent random variables, $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ does not hold.

Proof. Let $Y = X$, and X takes value 1 with probability 1/2, takes value −1 with the other probability 1/2. In this case, $\mathbb{E}[X] = \mathbb{E}[Y] = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$ but $\mathbb{E}[XY] = 1 \times \frac{1}{2} + 1 \times \frac{1}{2} = 1$. □

2.19 Prove Proposition 2.8.
Proposition 2.8. If $X \geq 0$ is a non-negative random variable, then

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > x)\, dx\,.$$

Hint Argue that $X(\omega) = \mathbb{1}\{[0, X(\omega)]\}(x)dx$ and exchange the integrals. Use the Fubini-Tonelli theorem to justify the exchange of integrals.

Proof. As the hint suggests, $X(\omega) = \int_{[0,\infty)} \mathbb{I}\{[0, X(\omega)]\}(x)dx$. Hence, we have

$$\begin{aligned}
\mathbb{E}[X(\omega)] &= \mathbb{E}\left[\int_{[0,\infty)} \mathbb{I}\{[0, X(\omega)]\}(x)dx\right] \\
&= \int_{[0,\infty)} \mathbb{E}\left[\mathbb{I}\{[0, X(\omega)]\}(x)\right] dx \\
&= \int_{[0,\infty)} P(X(\omega) > x)dx
\end{aligned} \tag{5}$$

where the second equality is given by Fubini–Tonell theorem.

□

2.20 Prove Theorem 1

Theorem 1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}$ be sub-σ-algebras of $\mathcal{F}$ and X, Y integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. The following hold true:*

1 *If $X \geq 0$, then $E[X|\mathcal{G}] \geq 0$ almost surely.*

2 *$E[1|\mathcal{G}] = 1$ almost surely.*

3 *$E[X + Y|\mathcal{G}] = E[X|\mathcal{G}] + E[Y|\mathcal{G}]$ almost surely.*

4 *$E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$ almost surely if $E[XY]$ exists and Y is $\mathcal{G}$-measurable.*

5 *If $\mathcal{G}_1 \subset \mathcal{G}_2$, then $E[X|\mathcal{G}_1] = E[E[X|\mathcal{G}_2]|\mathcal{G}_1]$*

6 *If $\sigma(X)$ is independent of $\mathcal{G}_2$ given $\mathcal{G}_1$, then $E[X|\sigma(\mathcal{G}_1 \bigcup \mathcal{G}_2)] = E[X|\mathcal{G}_1]$ almost surely.*

7 *If $\mathcal{G} = \{\emptyset, \Omega\}$ is the trival σ-algebra, then $E[X|\mathcal{G}] = E[X]$ almost surely.*

Proof. 1 If $P(E[X|\mathcal{G}] \geq 0) = p < 1$, choose $H = \{\omega | E[X|\mathcal{G}](\omega) < 0\} \in \mathcal{G}$. Based on the definition of conditional expectation,

$$0 \leq \int_H X dP = \int_H E[X|\mathcal{G}]dP < 0 \tag{6}$$

which causes contradiction.

2 For all $H \in \mathcal{G}$,

$$P(H) = \int_H 1dP = \int_H E[1|\mathcal{G}]dP \tag{7}$$

which means $E[1|\mathcal{G}] = 1$ a.s..

3 Based on the linearity, for all $H \in \mathcal{G}$

$$\begin{aligned}\int_H E[X+Y|\mathcal{G}]dP &= \int_H X + Y dP \\ &= \int_H XdP + \int_H YdP \\ &= \int_H E[X|\mathcal{G}]dP + \int_H E[Y|\mathcal{G}]dP \\ &= \int_H E[X|\mathcal{G}] + E[Y|\mathcal{G}]dP\end{aligned} \tag{8}$$

4 When Y is a simple function, $Y = \sum_{i=1}^n \alpha_i I\{A_i\}(A_i \in \mathcal{G}$ since Y is $\mathcal{G}$-measurable), for all $H \in \mathcal{G}$,

$$\begin{aligned}\int_H E[XY|\mathcal{G}]dP &= \int_H XYdP \\ &= \int_H \sum_{i=1}^n \alpha_i I\{A_i\}XdP \\ &= \sum_{i=1}^n \alpha_i \int_{H\bigcap A_i} XdP \\ &= \sum_{i=1}^n \alpha_i \int_{H\bigcap A_i} E[X|\mathcal{G}]dP \\ &= \int_H YE[X|\mathcal{G}]dP\end{aligned} \tag{9}$$

When Y is a non-negative random variable, $Y = \sup S = \sup\{h \geq 0 : h \text{ is simple and } h \leq Y\}$, let $X = X^+ - X^-$, where $X^+ = \max 0, X$. For all $H \in \mathcal{G}$,

$$\begin{aligned}\int_H E[XY|\mathcal{G}]dP &= \int_H (X^+ - X^-)YdP \\ &= \int_H X^+YdP - \int_H X^-YdP \\ &= \sup_{h\in S}\int_H hX^+dP - \sup_{h\in S}\int_H hX^-dP \text{ (Dominated convergence Thm)} \\ &= \int_H YE[X|\mathcal{G}]dP\end{aligned} \tag{10}$$

When Y is general random variable, $Y = Y^+ - Y^-$, using the linearity of integration, we can prove the statement.

5 For any $H \in \mathcal{G}_1 \subset \mathcal{G}_2$,

$$\int_H E[X|\mathcal{G}_1]dP = \int_H XdP = \int_H E[X|\mathcal{G}_2]dP = \int_H E[E[X|\mathcal{G}_2]|\mathcal{G}_1]dP \tag{11}$$

6

7 $E[X]$ is a constant, which is measurable on $\mathcal{G}$. We only need to prove that

$$\begin{cases}\int_\emptyset E[X|\mathcal{G}]dP = \int_\emptyset E[X]dP = 0 \\ \int_\Omega E[X|\mathcal{G}]dP = \int_\Omega E[X]dP = E[X]\end{cases} \tag{12}$$

□

Chapter 3 Stochastic Processes and Markov Chains

3.1 Fill in the details of Theorem 3.1:

(a) Prove that $F_t \in \{0, 1\}$ is a Bernoulli random variable for all $t \geq 1$.

(b) In what follows, equip $\mathcal{S}$ with $\mathbb{P} = \lambda$, the uniform probability measure. Show that for any $t \geq 1$, F_t is uniformly distributed: $\mathbb{P}(F_t = 1) = \mathbb{P}(F_t = 0) = 1/2$.

(c) Show that $(F_t)_{t=1}^{\infty}$ are independent.

(d) Show that $(X_{m,t})_{t=1}^{\infty}$ is an independent sequence of Bernoulli random variables that are uniformly distributed.

(e) Show that $X_t = \sum_{t=1}^{\infty} X_{m,t} 2^{-t}$ is uniformly distributed on $[0, 1]$

(f) Show that $(X_t)_{t=1}^{\infty}$ are independent.

Proof. (a) It can be shown that F_t has the following form:

$$F_t(x) = \mathbb{1}\{x \in U_t\}, U_t = \{1\} \cup \bigcup_{0 \leq s \leq 2^{t-1}} \left[\frac{2s-1}{2^t}, \frac{2s}{2^t}\right)$$

.

Since $U_t \in \mathcal{B}([0, 1])$, thus we can show F_t is $\mathcal{B}([0, 1])$-mearuable. And $F_t \in \{0, 1\}$, thus it is Bernoulli random variable.

(b) $\mathbb{P}(F_t = 1) = \mathbb{P}(U_t) = \sum_{s=0}^{2^{t-1}} \frac{1}{2^t} = \frac{2^{t-1}}{2^t} = \frac{1}{2}$. And $\mathbb{P}(F_t = 0) = 1 - \mathbb{P}(F_t = 1) = \frac{1}{2}$.

(c) Given an index set $K \subseteq \mathbb{N}+$ we need to show that $\{F_k, k \in K\}$ are independent. Or equivalently, that

$$\mathbb{P}\left(\bigcap_{k \in K}\right) = \prod_{k \in K} \mathbb{P}(U_k) = 2^{-|K|}.$$

Let $k = \max K$. Then

$$\lambda\left(U_k \cap \cup_{j \in K \setminus \{k\}} U_j\right) = \frac{1}{2} \lambda\left(\cup_{j \in K \setminus \{k\}} U_j\right).$$

Then by induction we can get the desired result.

(d) It follows directly from the definition of independence that any subsequence of an independent sequence is also an independent sequence. Thus from (b) we can get

$$\mathbb{P}(X_{m,t} = 0) = \mathbb{P}(X_{m,t} = 1) = \frac{1}{2}.$$

(e) By definition, $X_t = \sum_{t=1}^{\infty}$ is a weighted sum of an independent sequence of uniform Bernoully random variables. Thus X_t has the same property as random variable $Y = \sum_{t=1}^{\infty} F_t 2^{-t}$, and $Y(x) = x, \mathbb{P}(Y \le x) = x$ thus we can also conclude X_t is uniformly distributed.

(f) The disjoint subsets of independent random variables are also independent.

□

3.2 (Martingales and optional stopping) Let $(X_t)_{t=1}^{\infty}$ be an infinite sequence of independent Rademacher random variables and $S_t = \sum_{s=1}^{t} X_s 2^{s-1}$.

(a) Show that $(S_t)_{t=0}^{\infty}$ is a martingale.

(b) Let $\tau = \min\{t : S_t = 1\}$ and show that $\mathbb{P}(\tau < \infty) = 1$.

(c) What is $\mathbb{E}[S_\tau]$?

(d) Explain why this does not contradict Doob's optional stopping theorem.

Proof. (a) First of all, we can observe that S_t is $\mathcal{F}_t$-measurable given that it is actually a function of random variables $X_1, \cdots, X_t$ on $(\Omega, \mathcal{F}, \mathbb{P})$.

Next, we have

$$\begin{aligned}
\mathbb{E}[S_t \mid \mathcal{F}_{t-1}] &= \mathbb{E}\left[S_{t-1} + X_t 2^{t-1} \mid \mathcal{F}_{t-1}\right] \\
&= \mathbb{E}[S_{t-1} \mid \mathcal{F}_{t-1}] + \mathbb{E}\left[X_t 2^{t-1} \mid \mathcal{F}_{t-1}\right] \\
&= S_{t-1} + \mathbb{E}\left[X_t 2^{t-1} \mid \mathcal{F}_{t-1}\right] \\
&= S_{t-1} + \frac{1}{2}2^{t-1} + \frac{1}{2}(-2^{t-1}) \\
&= S_{t-1}.
\end{aligned}$$

Finally, as a weighted sum of finite random varialbes, S_t is surely integrable.

(b) Notice that

$$\begin{aligned}
\mathbb{P}(\tau = n) &= \mathbb{P}(\tau \neq 1)\mathbb{P}(\tau \neq 2 \mid \tau \neq 1) \cdots \mathbb{P}(\tau = n \mid \tau \notin \{1, \cdots n-1\}) \\
&= \mathbb{P}(X_1 = -1)\mathbb{P}(X_2 = -1 \mid X_1 = -1) \cdots \mathbb{P}(X_n = 1 \mid X_1 = \cdots = X_{n-1} = -1) \\
&= \frac{1}{2^n},
\end{aligned}$$

which implies $\mathbb{P}(\tau = \infty) = \lim_{n\to\infty} \mathbb{P}(\tau = n) = 0$ and hence $\mathbb{P}(\tau < \infty) = 1 - \mathbb{P}(\tau = \infty) = 1$.

(c) It is obvious that $\mathbb{E}(S_\tau) = 1$ according to our stopping rule.

(d) The result of (c) does not contradict Doob's optional stopping theorem because the conditions are not satisfied. Firstly, $\mathbb{P}(\tau > n) = \frac{1}{2^n} > 0$ for any $n \in \mathbb{N}$, which breaks the first condition. Next, $|S_{t+1} - S_t| = 2^t$ and $|S_{t\wedge\tau}| = |\sum_{s=1}^{t\wedge\tau} X_s 2^{s-1}$ can neither not be bounded, which breaks the second and third condition.

□

3.3 (Martingales and optional stopping (ii)) Give an example of a martingale $(S_n)_{n=0}^{\infty}$ and stopping time τ such that

$$\lim_{n\to\infty} \mathbb{E}[S_{\tau\wedge n}] \neq \mathbb{E}[S_\tau].$$

Proof. Notice that $(S_{\tau\wedge n})_n^{\infty}$ is also a martingale (which is usually called stopped martingale). This can be shown by formulating $S_{\tau\wedge n} = S_{\tau\wedge(n-1)} + \mathbb{I}\{\tau \le n\}(S_n - S_{n-1})$ and checking

$$\begin{aligned}
\mathbb{E}[S_{\tau\wedge n}|\mathcal{F}_{n-1}] &= \mathbb{E}[S_{\tau\wedge(n-1)} + \mathbb{I}\{\tau \le n\}(S_n - S_{n-1})|\mathcal{F}_{n-1}] \\
&= \mathbb{E}[S_{\tau\wedge(n-1)}|\mathcal{F}_{n-1}] + \mathbb{E}[\mathbb{I}\{\tau \le n\}(S_n - S_{n-1})|\mathcal{F}_{n-1}] \\
&= S_{\tau\wedge(n-1)} + \mathbb{I}\{\tau \le n\}(S_{n-1} - S_{n-1}) \\
&= S_{\tau\wedge(n-1)}.
\end{aligned}$$

Therefore, we have $\mathbb{E}(S_{\tau\wedge n}) = \mathbb{E}(S_{\tau\wedge 0}) = \mathbb{E}(S_0)$. Now it suffices to give the same example as in 3.2 that $\lim_{n\to\infty} \mathbb{E}\left[S_{\tau\wedge n}\right] = \mathbb{E}(S_0) \neq \mathbb{E}\left[S_\tau\right]$. $\square$

3.4 If $X_t \geq 0$ is dropped, $\mathbb{E}[X_\tau|\{\tau \leq n\}] \geq \mathbb{E}[\varepsilon|\{\tau \leq n\}]$ not always true.
3.5 Let $(\Omega, \mathcal{F})$ and $(\chi, \mathcal{G})$ be measurable spaces, $X : \chi \to \mathbb{R}$ be a random variable and $K : \Omega \times \mathcal{G} \to [0, 1]$ a prbability kernel from $(\Omega, \mathcal{F})$ to $(\chi, \mathcal{G})$. Define function $U : \Omega \to \mathbb{R}$ by $U(\omega) = \int_\chi X(x)K(\omega, dx)$ and assume that $U(\omega)$ exists for all ω. Prove that U is measurable.

Proof. When $X(x)$ is simple function, $X(x) = \sum_{i=1}^n \alpha_i I\{A_i\}$,

$$\begin{aligned} U(\omega) &= \int_\chi \sum_{i=1}^n \alpha_i I\{A_i\} K(w, dx) \\ &= \sum_{i=1}^n \alpha_i K(w, A_i \bigcap \chi) \text{ is measuable} \end{aligned} \tag{13}$$

When $X(x)$ is non-negative function, $X(x) = \sup\{h : h \text{ is simple and } 0 \leq h \leq X\}$,

$$\begin{aligned} U(\omega) &= \int_\chi \sup h(x)K(w, dx) \\ &= \sup \int_\chi h(x)K(w, dx) \text{ is measurable (due to dominated convergence thm)} \end{aligned} \tag{14}$$

When $X(x)$ is general function, $X = X^+ - X^-$, use the linearity of integration, we can still prove that U is measurable. $\square$

3.6(Limits of increasing stopping times are stopping times) Let $(\tau_n)_\infty$ be an almost surely increasing sequence of $\mathbb{F}$-stopping times on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathbb{F} = (\mathcal{F}_n)_{n=1}^\infty$, which means that $\tau_n(\omega) \leq \tau_{n+1}(\omega)$ for all $n \geq 1$ almost surely. Prove that $\tau(\omega) = \lim_{n\to\infty} \tau_n(\omega)$ is a $\mathbb{F}$-stopping time.

Proof. To show τ is a stopping time, we want to show that $\forall t, \mathbb{1}\{\tau \leq t\} \in \mathcal{F}_t$.

$$\mathbb{1}\{\tau \leq t\} = \mathbb{1}\left\{\lim_{n\to\infty} \tau_n \leq t\right\} = \mathbb{1}\left\{\sup_n \tau_n \leq t\right\} = \cap_n \mathbb{1}\{\tau_n \leq t\}$$

Since $\forall n, \tau_n$ is a stopping time, then $\mathbb{1}\{\tau_n \leq t\} \in \mathcal{F}_t, \forall t$. And $\forall t, \mathcal{F}_t$ is a σ-algebra. Thus $\forall t, \mathbb{1}\{\tau \leq t\} = \cap_n \mathbb{1}\{\tau_n \leq t\} \in \mathcal{F}_t$. $\square$

3.7 (Properties of stopping times) Let $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{N}}$ be a filtration, and τ, τ_1, τ_2 be stopping times with respect to $\mathbb{F}$. Show the following:

(a) $\mathcal{F}_\tau$ is a σ-algebra

(b) If $\tau = k$ for some $k \geq 1$, then $\mathcal{F}_\tau = \mathcal{F}_k$.

(c) If $\tau_1 \leq \tau_2$, then $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

(d) τ is $\mathcal{F}_\tau$-measurable.

(e) If (X_t) is $\mathbb{F}$-adapted, then X_τ is $\mathcal{F}_\tau$-measurable.

(f) $\mathcal{F}_\tau$ is the smallest σ-algebra such that all $\mathbb{F}$-adapted sequences (X_t) satisfy X_τ is $\mathcal{F}_\tau$-measurable.

Proof. (a) In this problem, we need to prove that $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t\}$. According to the definition of σ-algebra, here we just need to show each requirement is satisfied. First, the whole set is $\{\tau \leq t\}$, it is obvious that $\{\tau \leq t\} \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t$. Second, $\forall A \in \mathcal{F}_\tau, A^c \cap \{\tau \leq t\} = \{\tau \leq t\} - A \cap \{\tau \leq t\}$, since both $\{\tau \leq t\}$ and $A \cap \{\tau \leq t\}$ are in $\mathcal{F}_t, \forall t$, thus $A^c \in \mathcal{F}_t, \forall t$ holds. Last, $\forall A_i \in \mathcal{F}_\tau, i = 1, 2, ..., (\bigcup_i A_i) \cap \{\tau \leq t\} = \bigcup_i (A_i \cap \{\tau \leq t\})$, since $\mathcal{F}_t$ is a σ-algebra and $A_i \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t$, we can conclude $(\bigcup_i A_i) \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t$. Above all, we have shown that $\mathcal{F}_\tau$ is a σ-algebra on the set $\{\tau \leq t\}$.

(b) We first prove $\mathcal{F}_\tau \subseteq \mathcal{F}_k$, which means $\forall A \in \mathcal{F}_\tau, A \in \mathcal{F}_k$. According to the definition, $\forall A \in \mathcal{F}_\tau, A \cap \{\tau \leq k\} \in \mathcal{F}_k$, and $\{\tau \leq k\}$ holds a.s. as $\tau = k$. Thus $A = A \cap \{\tau \leq k\} \in \mathcal{F}_k$.

We then prove $\mathcal{F}_k \subseteq \mathcal{F}_\tau$, that is $\forall B \in \mathcal{F}_k, B \in \mathcal{F}_\tau$. When $t \geq k$, $B \cap \{\tau \leq t\} = B$, and since $B \in \mathcal{F}_k \subseteq \mathcal{F}_t$, we could have $B \cap \{\tau \leq t\} \in \mathcal{F}_t$. And for $t < k$, $B \cap \{\tau \leq t\} = \emptyset \in \mathcal{F}_t$. Thus we could conclude that $B \in \mathcal{F}_\tau$.

(c) We want to prove $\forall A \in \mathcal{F}_{\tau_1}$, it holds that $A \in \mathcal{F}_{\tau_2}$. Recall that

$$\begin{aligned}\mathcal{F}_{\tau_1} &= \{A \in \mathcal{F}_\infty : A \cap \{\tau_1 \leq t\} \in \mathcal{F}_t, \forall t\} \ ,\\ \mathcal{F}_{\tau_2} &= \{A \in \mathcal{F}_\infty : A \cap \{\tau_2 \leq t\} \in \mathcal{F}_t, \forall t\} \ .\end{aligned}$$

Since $\tau_1 \leq \tau_2$, $\{\tau_2 \leq t\} \subseteq \{\tau_1 \leq t\}, \forall t$. According to the definition, we have $\forall A \in \mathcal{F}_{\tau_1}$, $A \cap \{\tau_1 \leq t\} \in \mathcal{F}_t, \forall t$. Then $A \cap \{\tau_2 \leq t\} = A \cap \{\tau_1 \leq t\} - (\{\tau_1 \leq t\} - \{\tau_2 \leq t\})$. And all of $A \cap \{\tau_1 \leq t\}$, $\{\tau_1 \leq t\}$, $\{\tau_2 \leq t\}$ are in $\mathcal{F}_t, \forall t$, we then have $A \cap \{\tau_2 \leq t\} \in \mathcal{F}_t, \forall t$. Thus there is $A \in \mathcal{F}_{\tau_2}$ and further $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

(d)

(e) We first clearly present the definition of $X_\tau = X_{\tau(\omega)}(\omega)$. If $\tau(\omega) = t$ for some t, then $X_{\tau(\omega)}(\omega) = X_t(\omega)$. If $\tau(\omega) = \infty$, then $X_{\tau(\omega)}(\omega) = \lim_{t\to\infty} X_t(\omega)$. To prove X_τ is $\mathcal{F}_\tau$-measurable, we want to prove for any constant $a \in \mathbb{R}$, $X_\tau^{-1}((-\infty, a)) = \{\omega : X_\tau(\omega) < a\} \in \mathcal{F}_\tau$, where $\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t\}$. We first prove $\{\omega : X_\tau(\omega) < a\} \in \mathcal{F}_\infty$, it can be written as

$$\{\omega : X_\tau(\omega) < a\} = \left\{\omega : \bigcup_{t\in\mathbb{N}} \{\tau(\omega) = t\} \cap \{X_t(\omega) < a\} \bigcup \{\tau(\omega) = \infty\} \cap \left\{\lim_{t\to\infty} X_t(\omega) < a\right\}\right\}$$

Since $\forall t$, $\{\tau(\omega) = t\} \cap \{X_t(\omega) < a\} \in \mathcal{F}_t$, thus $\bigcup_{t\in\mathbb{N}} \{\tau(\omega) = t\} \cap \{X_t(\omega) < a\} \in \mathcal{F}_\infty$. And $\{\tau(\omega) = \infty\} = \bigcap_{t\in\mathbb{N}} \{\tau(\omega) > t\} \in \mathcal{F}_\infty$. For the last part,

$$\left\{\lim_{t\to\infty} X_t(\omega) < a\right\} = \cap_{n>0} \cup_{N\in\mathbb{N}} \cap_{t,s>N} \left\{|X_t(\omega) - X_s(\omega)| \leq \frac{1}{2^n}\right\} \bigcap \cup_{N\in\mathbb{N}} \cap_{t>N} \{X_t(\omega) < a\} \ ,$$

where the first part represents the existence of the limitation and the second part represents the limitation is less than a. It is obvious that $\left\{|X_t(\omega) - X_s(\omega)| \leq \frac{1}{2^n}\right\} \in \mathcal{F}_{\max\{t,s\}}$, and thus $\cup_{N\in\mathbb{N}} \cap_{t,s>N} \left\{|X_t(\omega) - X_s(\omega)| \leq \frac{1}{2^n}\right\} \in \mathcal{F}_\infty$. Similarly the second part $\cup_{N\in\mathbb{N}} \cap_{t>N} \{X_t(\omega) < a\} \in \mathcal{F}_\infty$. Above all, we have proved $\{\omega : X_\tau(\omega) < a\} \in \mathcal{F}_\infty$. Next we will show $\forall t, \{X_\tau(\omega) < a\} \cap \{\tau(\omega) \leq t\} \in \mathcal{F}_t$.

$$\{\omega : X_\tau(\omega) < a\} \cap \{\tau(\omega) \leq t\} = \bigcup_{s\leq t} \{\tau(\omega) = s\} \cap \{X_s(\omega) < a\} \tag{15}$$

As $\{\tau(\omega) = s\} \cap \{X_s(\omega) < a\} \in \mathcal{F}_s$, we have $\bigcup_{s\leq t} \{\tau(\omega) = s\} \cap \{X_s(\omega) < a\} \in \mathcal{F}_t, \forall t$.

Above all, we have proved for any constant $a \in \mathbb{R}$, $X_\tau^{-1}((-\infty, a)) = \{\omega : X_\tau(\omega) < a\} \in \mathcal{F}_\tau$

(f)

□

3.8 Prove Theorem 3.10.
Theorem 3.10 Let $(X_t)_{t=0}^n$ be a submartingale with $X_t \geq 0$ almost surely for all t. Then for any $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{t\in\{0,1,\ldots,n\}} X_t \geq \varepsilon\right) \leq \frac{\mathbb{E}[X_n]}{\varepsilon} \ .$$

Proof. Define event $A = \left\{\max_{t\in\{0,1,\ldots,n\}} X_t < \varepsilon\right\}$, $\forall t = 0, .., n, A_t = \{\forall i < t : X_i < \varepsilon, X_t \geq \varepsilon\}$. Note that event A and $A_t, \forall t$ are exclusive.

According to the property of submartingale and $X_t \geq 0, \forall t$, we have

$$X_t \mathbb{1}\{A_t\} \leq \mathbb{E}[X_n \mid \mathcal{F}_t] \mathbb{1}\{A_t\} = \mathbb{E}[X_n \mathbb{1}\{A_t\} \mid \mathcal{F}_t] \ ,$$

further according to the tower rule, we have

$$\mathbb{E}\left[X_t \mathbb{1}\{A_t\}\right] \leq \mathbb{E}\left[\mathbb{E}\left[X_n \mathbb{1}\{A_t\} \mid \mathcal{F}_t\right]\right] = \mathbb{E}\left[X_n \mathbb{1}\{A_t\}\right] .$$

Above all, there is

$$\begin{aligned}
\mathbb{E}\left[X_n\right] &= \mathbb{E}\left[X_n \mathbb{1}\{A\}\right] + \sum_{t=0}^{n} \mathbb{E}\left[X_n \mathbb{1}\{A_t\}\right] \\
&\geq \sum_{t=0}^{n} \mathbb{E}\left[X_n \mathbb{1}\{A_t\}\right] \\
&\geq \sum_{t=0}^{n} \mathbb{E}\left[X_t \mathbb{1}\{A_t\}\right] \\
&\geq \varepsilon \cdot \sum_{t=0}^{n} \mathbb{E}\left[\mathbb{1}\{A_t\}\right] \\
&= \varepsilon \cdot \sum_{t=0}^{n} \mathbb{P}\left(A_t\right) \\
&= \varepsilon \cdot \mathbb{P}\left(\bigcup_{t=0}^{n} A_t\right) \\
&= \varepsilon \cdot \mathbb{P}\left(\max_{t \in \{0,1,\ldots,n\}} X_t \geq \varepsilon\right)
\end{aligned}$$

$\square$

Chapter 4 Stochastic Bandits

4.1 By definition

$$\begin{aligned}
R_n(\pi, v) &= n\mu^*(v) - \mathbb{E}[\sum_{t=1}^{n} X_t] \\
&= \sum_{t=1}^{n} \mu^*(v) - \sum_{t=1}^{n} \mathbb{E}[X_t] \\
&= \sum_{t=1}^{n} [\mu^* - \mu_{A_t}]
\end{aligned}$$

(a) $\mu^* = \max \mu_a \geq \mu_{A_t} \Rightarrow R_n(\pi, v) = \sum_{t=1}^{n}[\mu^* - \mu_{A_t}] \geq 0$.

(b) If π choose $A_t \in \arg\max_a \mu_a$ for all $t \in [n] \Rightarrow \sum_{t=1}^{n}[\mu^* - \mu_{A_t}] = 0$.

(c) If $R_n(\pi, v) = 0$ for some policy π ,then $A_t \in \arg\max_a \mu_a \Rightarrow \mathbb{P}(\mu_{A_t} = \mu^*) = 1$.

4.2(Uniqueness of law) Prove Proposition 4.8.

Proposition 1. *Suppose that $\mathbb{P}$ and $\mathbb{Q}$ are probability measures on an arbitrary measurable space $(\Omega, \mathcal{F})$ and $A_1, X_1, ..., A_n, X_n$ are random variables on Ω, where $A_t \in [k]$ and $X_t \in \mathbb{R}$. If both $\mathbb{P}$ and $\mathbb{Q}$ satisfy (a) and (b), then the law of the outcome under $\mathbb{P}$ is the same as under $\mathbb{Q}$:*

$$\mathbb{P}_{A_1, X_1, \ldots, A_n, X_n} = \mathbb{Q}_{A_1, X_1, \ldots, A_n, X_n}\,.$$

Recall the condition:

(a) the conditional distribution of action A_t given $A_1, X_1, ..., A_{t-1}, X_{t-1}$ is $\pi_t(\cdot \mid A_1, X_1, ..., A_{t-1}, X_{t-1})$ almost surely.

(b) the conditional distribution of reward X_t given $A_1, X_1, ..., A_t$ is P_{A_t} almost surely.

Proof.

$$\begin{aligned}
\mathbb{P}_{A_1, X_1, \ldots, A_n, X_n} &= \mathbb{P}(A_1)\mathbb{P}(X_1 \mid A_1) \ldots \mathbb{P}(A_n \mid A_1, X_1, ..., A_{n-1}, X_{n-1})\mathbb{P}(X_n \mid A_1, X_1, ..., A_{n-1}, X_{n-1}, A_n) \\
&= \pi_1(A_1)P_{A_1} \ldots \pi_n(A_n \mid A_1, X_1, ..., A_{n-1}, X_{n-1})P_{A_n} \\
&= \mathbb{Q}(A_1)\mathbb{Q}(X_1 \mid A_1) \ldots \mathbb{Q}(A_n \mid A_1, X_1, ..., A_{n-1}, X_{n-1})\mathbb{Q}(X_n \mid A_1, X_1, ..., A_{n-1}, X_{n-1}, A_n) \\
&= \mathbb{Q}_{A_1, X_1, \ldots, A_n, X_n}\,.
\end{aligned}$$

□

4.3 Denote $h_t = a_1, x_1, \ldots, a_t, x_t$.

(a) According to the definition of conditional probability and marginal distribution, we have

$$\begin{aligned}
p_{v\pi}(a_n \mid h_{n-1}) &= \frac{p_{v\pi}(h_{n-1}, a_n)}{p_{v\pi}(h_{n-1})} \\
&= \frac{\int_{\mathbb{R}} p_{v\pi}(h_n) dx_n}{p_{v\pi}(h_{n-1})} \\
&= \frac{\int_{\mathbb{R}} \prod_{t=1}^{n} \pi\left(a_t \mid h_{t-1}\right) p_{a_t}\left(x_t\right) dx_n}{p_{v\pi}(h_{n-1})} \\
&= \frac{\prod_{t=1}^{n-1} \pi\left(a_t \mid h_{t-1}\right) p_{a_t}\left(x_t\right)}{p_{v\pi}(h_{n-1})} \int_{\mathbb{R}} \pi\left(a_n \mid h_{n-1}\right) p_{a_n}\left(x_n\right) dx_n \\
&= \pi\left(a_n \mid h_{n-1}\right) \int_{\mathbb{R}} p_{a_n}\left(x_n\right) dx_n \\
&= \pi\left(a_n \mid h_{n-1}\right)
\end{aligned}$$

(b) According to the definition of conditional probability and marginal distribution, we have

$$\begin{aligned}
p_{v\pi}(x_n \mid h_{n-1}, a_n) &= \frac{p_{v\pi}(h_n)}{p_{v\pi}(h_{n-1}, a_n)} \\
&= \frac{p_{v\pi}(h_n)}{\int_{\mathbb{R}} p_{v\pi}(h_n) dx_n} \\
&= \frac{p_{v\pi}(h_n)}{\int_{\mathbb{R}} \left[\prod_{t=1}^{n} \pi\left(a_t \mid h_{t-1}\right) p_{a_t}\left(x_t\right)\right] dx_n} \\
&= \frac{p_{v\pi}(h_n)}{\prod_{t=1}^{n-1} \pi\left(a_t \mid h_{t-1}\right) p_{a_t}\left(x_t\right)} \frac{1}{\int_{\mathbb{R}} \pi\left(a_n \mid h_{n-1}\right) p_{a_n}\left(x_n\right) dx_n} \\
&= \pi\left(a_n \mid h_{n-1}\right) p_{a_n}\left(x_n\right) \frac{1}{\pi\left(a_n \mid h_{n-1}\right)} \\
&= p_{a_n}\left(x_n\right)
\end{aligned}$$

4.4 Denote $h_t = a_1, x_1, \ldots, a_t, x_t$. The policy that mixes the policies can be defined as

$$\pi_t^{\circ}\left(a_t \mid h_{t-1}\right) = \frac{\sum_{\pi \in \Pi} p(\pi) \prod_{s=1}^{t} \pi_s\left(a_s \mid h_{s-1}\right)}{\sum_{\pi \in \Pi} p(\pi) \prod_{s=1}^{t-1} \pi_s\left(a_s \mid h_{s-1}\right)}$$

.

By the definition of the canonical probability space and the product of probability kernels,

$$\begin{aligned}
\mathbb{P}_{v\pi^{\circ}}(B) &= \sum_{a_1=1}^{k} \int_{\mathbb{R}} \cdots \sum_{a_n=1}^{k} \int_{\mathbb{R}} \mathbb{I}_B\left(h_n\right) v_{a_n}\left(dx_n\right) \pi_n^{\circ}\left(a_n \mid h_{n-1}\right) \cdots v_{a_1}\left(dx_1\right) \pi_1^{\circ}\left(a_1\right) \\
&= \sum_{\pi \in \Pi} p(\pi) \sum_{a_1=1}^{k} \int_{\mathbb{R}} \cdots \sum_{a_n=1}^{k} \int_{\mathbb{R}} \mathbb{I}_B\left(h_n\right) v_{a_n}\left(dx_n\right) \pi_n\left(a_n \mid h_{n-1}\right) \cdots v_{a_1}\left(dx_1\right) \pi_1\left(a_1\right) \\
&= \sum_{\pi \in \Pi} p(\pi) \mathbb{P}_{v\pi}(B),
\end{aligned}$$

where the second equality follows by substituting the definition of π_n° and induction.

Chapter 5 Concentration of Measure

5.1 (Variance of average) Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with mean μ and variance $\sigma^2 < \infty$. Let $\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} X_t$ and show that $\mathbb{V}[\hat{\mu}] = \mathbb{E}\left[(\hat{\mu}-\mu)^2\right] = \sigma^2/n$.

Proof. We start straight from the definition:

$$
\begin{aligned}
&\mathbb{V}[\hat{\mu}] \\
=&\mathbb{E}((\hat{\mu}-\mu)^2) \\
=&\mathbb{E}((\frac{1}{n}\sum_{t=1}^{n} X_t - \mu)^2) \\
=&\mathbb{E}(\frac{1}{n^2}\sum_{t=1}^{n}(X_t-\mu)^2) \\
=&\frac{1}{n^2}\sum_{t=1}^{n}\mathbb{E}(X_t-\mu)^2 \\
=&\frac{1}{n^2}\sum_{t=1}^{n}\sigma^2 \\
=&\frac{\sigma^2}{n}
\end{aligned}
$$

□

5.2 (Markov's inequality) Prove Markov's inequality (Lemma 5.1).

Lemma 1. *For any random variable X and $\epsilon > 0$, the following holds:*

$$\mathbb{P}\left(|X| \geq \epsilon\right) \leq \frac{\mathbb{E}\left[|X|\right]}{\epsilon}$$

Proof. Suppose, for a contradiction, that $\mathbb{P}\left(|X| \geq \epsilon\right) > \frac{\mathbb{E}[|X|]}{\epsilon}$, then by definition the expectation of $absX$ is

$$\mathbb{E}\left[|X|\right] = \int_0^\infty \mathbb{P}\left(|X| = x\right) x dx \geq \mathbb{P}\left(|X| \geq \epsilon\right)\epsilon > \mathbb{E}\left[|X|\right]$$

which contradicts the assumption. □

5.3 Compare the Gaussian tail probability bound on the right-hand side of (5.4) and the one on (5.2). What values of ε make one smaller than the other? Discuss your findings.

Proof. Two-slided tail directly in terms of the variance in (5.2) is $\mathbb{P}(|\hat{\mu}-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$. And in (5.4)

$$
\begin{aligned}
\mathbb{P}(\hat{\mu} \geq \mu + \varepsilon) &= \mathbb{P}\left(S_n/\sqrt{\sigma^2 n} \geq \varepsilon\sqrt{n/\sigma^2}\right) \approx \mathbb{P}\left(Z \geq \varepsilon\sqrt{n/\sigma^2}\right) \\
&\leq \sqrt{\frac{\sigma^2}{2\pi n\varepsilon^2}} \exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right)
\end{aligned}
$$

We need to compare $\frac{\sigma^2}{2n\varepsilon^2}$ and $\sqrt{\frac{\sigma^2}{2\pi n\varepsilon^2}}\exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right)$ and we can define:

$$f(\varepsilon)=\frac{\sigma^2}{2n\varepsilon^2}-\sqrt{\frac{\sigma^2}{2\pi n\varepsilon^2}}\exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right)$$

Let $x=\frac{\sigma^2}{2n\varepsilon^2}$, then

$$f(x)=x-\frac{\sqrt{x}}{\sqrt{\pi}}\exp(\frac{1}{4x})$$

When $f(x)=0$, the Lambert form of the equation could be writen as:

$$\frac{1}{\pi}=x\exp(-\frac{1}{2x})$$

with the $\frac{1}{2x}=u$, the equation is transformed into

$$u\exp(u)=\frac{\pi}{2}$$

Solve the equation, $u=W_0(\frac{\pi}{2})$, and then $f(x)=0$ when

$$\varepsilon=\sqrt{\frac{W_0(\frac{\pi}{2})\sigma^2}{n}}$$

In order to judge the positive or negative of the $f(x)$, we need to judge the monotonicity in addition to solving the zero point. $f'(x)=x-\frac{\sqrt{x}}{\pi}exp(\frac{1}{4x})$, and from the picture, $f'(x)$ is always bigger than 0. Thus, when $\varepsilon>\sqrt{\frac{W_0(\frac{\pi}{2})\sigma^2}{n}}$, (5.2) is greater than (5.4), otherwise (5.2) is smaller.

$\square$

5.4 Let X be a random variable on $\mathbb{R}$ with density with respect to the Lebesgue measure of $p(x)=|x|\exp\left(-x^2/2\right)/2$. Show the following:

(a) $\mathbb{P}(|X|\geq\varepsilon)=\exp(-\varepsilon^2/2)$

(b) X is not $\sqrt{2-\varepsilon}$-subgaussian for any $\varepsilon>0$.

Proof. (a)

$$\begin{aligned}\mathbb{P}(|X|\geq\varepsilon)&=\int_{-\infty}^{-\varepsilon}\frac{|x|}{2}\exp\left(-x^2/2\right)dx+\int_{\varepsilon}^{+\infty}\frac{|x|}{2}\exp\left(-x^2/2\right)dx\\&=\int_{-\infty}^{-\varepsilon}-\frac{x}{2}\exp\left(-x^2/2\right)/2dx+\int_{\varepsilon}^{\infty}\frac{x}{2}\exp\left(-x^2/2\right)dx\\&=\int_{\varepsilon}^{\infty}x\exp\left(-x^2/2\right)dx\\&=\exp\left(-\varepsilon^2/2\right).\end{aligned}$$

(b) According to the definition of the sub-gaussian random variables, to prove X is not $\sqrt{2-\varepsilon}$-subgaussian, we want to prove $\exists\lambda\in\mathbb{R}$ such that

$$\mathbb{E}\left[\exp\left(\lambda X\right)\right]\leq\exp\left(\frac{\lambda^2(2-\varepsilon)}{2}\right)\triangleq g(\lambda).$$

For the LHS, there is

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\lambda X\right)\right] &= \int_{-\infty}^{+\infty} \exp(\lambda x)\frac{|x|}{2}\exp(-x^2/2)dx \\
&= \frac{1}{2}\exp(\lambda^2/2)\int_{-\infty}^{+\infty} |t+\lambda|\exp(-t^2/2)dt \\
&= \frac{\exp(\frac{\lambda^2}{2})}{2}\left(\int_{-\infty}^{-\lambda}(-t-\lambda)\exp(-t^2/2)dt + \int_{-\lambda}^{+\lambda}(t+\lambda)\exp(-t^2/2)dt + \int_{\lambda}^{+\infty}(t+\lambda)\exp(-t^2/2)dt\right) \\
&= \frac{\exp(\frac{\lambda^2}{2})}{2}\left(2\exp(-\frac{\lambda^2}{2}) + 2\lambda\int_0^{\lambda}\exp(-t^2/2)dt\right) \\
&= 1+\lambda\exp(\frac{\lambda^2}{2})\int_0^{\lambda}\exp(-t^2/2)dt \triangleq f(\lambda)\,,.
\end{aligned}
$$

Define $F(\lambda) = f(\lambda) - g(\lambda)$, now we want to prove $\exists\lambda \in \mathbb{R}, F(\lambda) > 0$. By computation, we found that $F(0) = 0, F'(0) = 0, F''(0) = \varepsilon > 0$. We can then conclude that $(0,0)$ is a minima of F, thus $\exists\delta > 0, \forall\lambda \in (-\delta,\delta), F(\lambda) > 0$. We then get the desired result.

□

5.5 (Berry–Esseen inequality) Let $X_1, X_2, \cdots, X_n$ be a sequence of independent and identically distributed random variables with mean μ, variance σ^2 and bounded third absolute moment:

$$\rho = \mathbb{E}[|X_1-\mu|^3] < \infty$$

Let $S_n = \sum_{t=1}^n (X_t-\mu)/\sigma$. The Berry–Esseen theorem shows that

$$\sup_x \left|\mathbb{P}\left(\frac{Sn}{\sqrt{n}} \le x\right) - \underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-y^2/2)dy}_{\Phi(x)}\right| \le \frac{C\rho}{\sqrt{n}}\,,$$

where $C < 1/2$ is a universal constant.

(a) Let $\hat{\mu}_n = \frac{1}{n}\sum_{t=1}^n X_t$ and derive a tail bound from the Berry–Esseen theorem. That is, give a bound of the form $\mathbb{P}(\hat{\mu}_n \ge \mu+\epsilon)$ for positive values of ϵ.

(b) Compare your bound with the one that can be obtained from the Cramér–Chernoff method. Argue pro- and contra- for the superiority of one over the other.

Proof. (a)

$$
\begin{aligned}
\mathbb{P}(\hat{\mu}_n \ge \mu+\epsilon) &= \mathbb{P}\left(\frac{1}{n}\sum_{t=1}^{n}(X_t-\mu) \ge \epsilon\right) \\
&= \mathbb{P}\left(\frac{\sum_{t=1}^n (X_t-\mu)}{\sqrt{n}\sigma} \ge \frac{\epsilon\sqrt{n}}{\sigma}\right) \\
&= 1-\mathbb{P}\left(\frac{\sum_{t=1}^n (X_t-\mu)}{\sqrt{n}\sigma} \le \frac{\epsilon\sqrt{n}}{\sigma}\right) \\
&\le 1-\Phi\left(\frac{\epsilon\sqrt{n}}{\sigma}\right) + \frac{C\rho}{\sqrt{n}} \\
&= \mathbb{P}\left(Z > \frac{\epsilon\sqrt{n}}{\sigma}\right) + \frac{C\rho}{\sqrt{n}} \quad \text{where} \quad Z \sim \mathcal{N}(0,1) \\
&\le \sqrt{\frac{\sigma^2}{2\pi n\epsilon^2}}\exp\left(-\frac{n\epsilon^2}{2\sigma^2}\right) + \frac{C\rho}{\sqrt{n}}\,.
\end{aligned} \tag{16}
$$

(b) We further assume that X_t is σ-subgaussian for all $t \in [n]$. Applying Cramér–Chernoff method shows that

$$\mathbb{P}(\hat{\mu} - \mu > \epsilon) \leq \exp\left(-\frac{n\epsilon^2}{2\sigma^2}\right). \tag{17}$$

We now discuss the pros and cons based on the additional multiplicative term $\sqrt{\frac{\sigma^2}{2\pi n\epsilon^2}}$ and the additive term $\frac{C\rho}{\sqrt{n}}$ in Eq.(16). When n is large enough, the terms $\sqrt{\frac{\sigma^2}{2\pi n\epsilon^2}}$ and $\frac{C\rho}{\sqrt{n}}$ will both be small. It follows that Eq.(16) will be tighter when $n \to \infty$ but will be looser when n is small than Eq.(17).

□

5.6 (CENTRAL LIMIT THEOREM) We mentioned that invoking the CLT to approximate the distribution of sums of independent Bernoulli random variables using a Gaussian can be a bad idea. Let $X_1, \ldots, X_n \sim \mathcal{B}(p)$ be independent Bernoulli random variables with common mean $p = p_n = \lambda/n$, where $\lambda \in (0, 1)$. For $x \in \mathbb{N}$ natural number, let $P_n(x) = \mathbb{P}(X_1 + \cdots + X_n = x)$.

(a) Show that $\lim_{n\to\infty} P_n(x) = e^{-\lambda}\lambda^x/(x!)$, which is a Poisson distribution with parameter λ.

(b) Explain why this does not contradict the CLT, and discuss the implications of the Berry-Esseen.

(c) In what way does this show that the CLT is indeed a poor approximation in some cases?

(d) Based on Monte Carlo simulations, plot the distribution of $X_1 + \cdots + X_n$ for $n = 30$ and some well-chosen values of λ. Compare the distribution to what you would get from the CLT. What can you conclude?

Proof. (a)

$$\begin{aligned}
P(X = x) &= l_{n\to\infty} \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x} \\
&= l_{n\to\infty} \begin{pmatrix} n \\ x \end{pmatrix} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= l_{n\to\infty} \frac{n!}{(n-x)!x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \\
&= l_{n=\infty} \frac{n!}{(n-x)!} \frac{1}{n^x} \frac{\lambda^x}{x!} e^{-\lambda} \quad 1
\end{aligned}$$

Then, we just need to proof the coefficient of possion is 1.

$$\ell_{h=\infty} \underbrace{\frac{n!}{(n-x)!} \frac{1}{n^x}}_{h\to\infty} = 1.$$

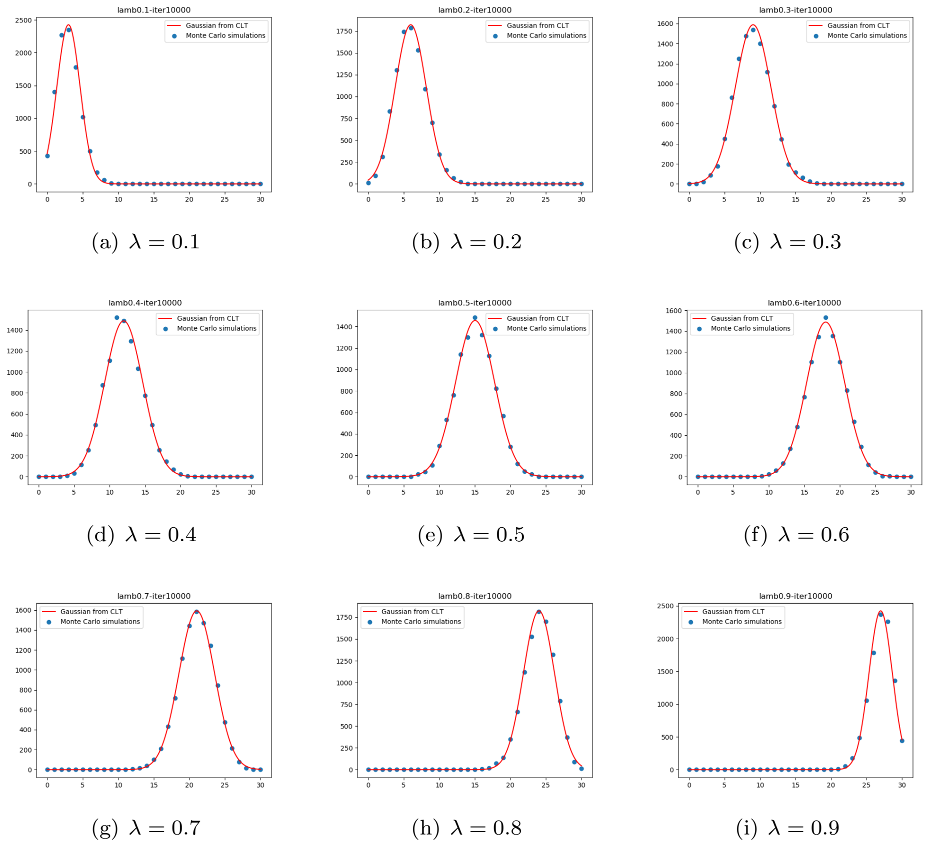
$$\frac{n(n-1)(n-2)\cdots(n-x+1)}{n} = 1.$$

Thus,

$$\lim_{n\to\infty} P_n(x) = e^{-\lambda}\lambda^x/(x!)$$

(b) According to the definition in section 5.2, central limit theorem is the limiting distribution of $S_n/\sqrt{n\sigma^2}$ as $n \to \infty$ is a Gaussian with mean zero and unit variance with $S_n = \sum_{t=1}^n (X_t - \mu)$.

But the conclusion in (a) is not conflict with central limit theorem, since $X_1, \ldots, X_n$ in Exp 5.6 with common mean $p = p_n = \lambda/n$ which does not satisfy the assumptions of variable's independence in the central limit theorem.



(a) $\lambda = 0.1$ (b) $\lambda = 0.2$ (c) $\lambda = 0.3$

(d) $\lambda = 0.4$ (e) $\lambda = 0.5$ (f) $\lambda = 0.6$

(g) $\lambda = 0.7$ (h) $\lambda = 0.8$ (i) $\lambda = 0.9$

Figure 1: Monte Carlo simulation and CLT results in 5.6.d

(d) The comparison of Monte Carlo simulation and CLT results can be found in Fig. 1. Our observation is that, when n becomes larger and can be regarded as continuous, the poisson distribution is more close to the normal distribution.

$\square$

5.7 (Properties of subgaussian random variables (i)) Prove Lemma 5.4.
Hint Use Taylor series.

Proof. (a) We consider using the Taylor series of the moment-generating function of r.v. X to prove the properties of the first moment and the second central moment of X. Specifically,

$$\mathbb{E}[\exp(tx)] = \mathbb{E}[\sum_{n=0}^{\infty} \frac{t^n}{n!} x^n] = \sum_{n=0}^{\infty} t^n \frac{\mathbb{E}[x^n]}{n!} \leq \exp(\frac{\sigma^2}{2} t^2) = \sum_{n=0}^{\infty} \frac{1}{n!} (\frac{\sigma^2}{2} t^2)^n ,$$

where the inequality is because X is a σ-subgaussian r.v. When $n = 2$,

$$1 + t\mathbb{E}[x] + \frac{t^2}{2}\mathbb{E}[x^2] \leq 1 + \frac{\sigma^2 t^2}{2} + g(t) , \tag{18}$$

where $g(t)$ satisfies that $\frac{g(t)}{t^2} \to 0$ for $t \to 0$. Based on this, one can see that $\mathbb{E}[x] \leq 0$ holds when $t \to 0_+$ and $\mathbb{E}[x] \geq 0$ holds when $t \to 0_-$, which implies that $\mathbb{E}[x] = 0$. Dividing both sides of Eq. (18) by t^2 shows that $\mathbb{E}[X^2] \leq \sigma^2$ and $\mathbb{V}[X^2] \leq \mathbb{E}[X^2] \leq \sigma^2$.

(b) $\forall C \in \mathbb{R}$, it holds that

$$\mathbb{E}[\exp(t \cdot cX)] \leq \exp\left(\frac{\sigma^2 \cdot (tc)^2}{2}\right) = \exp\left(\frac{(c\sigma)^2}{2} \cdot t^2\right) ,$$

which shows that cX is $|c|\sigma$-subgaussian immediately.

(c) One can see that

$$\begin{aligned}\mathbb{E}[\exp(t\cdot(X_1+X_2))] =& \mathbb{E}[\exp(tX_1)\cdot\exp(tX_2)] = \mathbb{E}[\exp(tX_1)]\cdot\mathbb{E}[\exp(tX_2)] \\ \leq& \exp(\frac{\sigma_1^2}{2}t^2)\cdot\exp(\frac{\sigma_2^2}{2}t^2) = \exp(\frac{\sigma_1^2+\sigma_2^2}{2}t^2)\,,\end{aligned}$$

which shows that X_1+X_2 is $\sqrt{\sigma_1^2+\sigma_2^2}$-subgaussian.

□

5.8 (Properties of subgaussian random variables (ii)) Let X_i be σ_i-subgaussian for $i\in\{1,2\}$ with $\sigma_i\geq 0$. Prove that X_1+X_2 is $(\sigma_1+\sigma_2)$-subgaussian. Do *not* assume independence of X_1 and X_2.

Proof. We start straight from the definition:

$$\begin{aligned}&\mathbb{E}[\exp(\lambda(X_1+X_2))]\\ \leq&\mathbb{E}[\exp(\lambda pX_1)]^{\frac{1}{p}}\mathbb{E}[\exp(\lambda qX_2)]^{\frac{1}{q}}\\ \leq&\exp(\lambda^2p^2\sigma_1^2/2)^{\frac{1}{p}}\exp(\lambda^2q^2\sigma_2^2/2)^{\frac{1}{q}}\\ =&\exp(\frac{\lambda^2(p\sigma_1^2+q\sigma_2^2)}{2})\\ =&\exp(\frac{\lambda^2(\sigma_1^2+\sigma_2^2)}{2}),\end{aligned}$$

where the first inequality holds according to Hölder's inequality and the last equality holds with $p=\frac{\sigma_1^2+\sigma_2^2}{\sigma_2^2}$.

□

5.9 (Properties of moment/cumulative-generating functions) Let X be a real-valued random variable and let $M_X(\lambda)=\mathbb{E}\left[\exp(\lambda X)\right]$ be its moment-generating function defined over $\mathrm{dom}(M_X)\subseteq\mathbb{R}$, where the expectation takes on finite values. Show that the following properties hold:

(a) M_X is convex, and in particular $\mathrm{dom}(M_X)$ is an interval containing zero.

(b) $M_X(\lambda)\geq e^{\lambda\mathbb{E}[X]}$ for all $\lambda\in\mathrm{dom}(M_X)$.

(c) For any λ in the interior of $\mathrm{dom}(M_X)$, M_X is infinitely many times differentiable.

(d) Let $M_X^{(k)}(\lambda)=\frac{d^k}{d\lambda^k}M_X(\lambda)$. Then, for λ in the interior of $\mathrm{dom}(M_X)$, $M^{(k)(\lambda)}=\mathbb{E}\left[X^k\exp(\lambda X)\right]$.

(e) Assuming 0 is in the interior of $\mathrm{dom}(M_X)$, $M^{(k)(0)}=\mathbb{E}\left[X^k\right]$ (hence the name of M_X).

(f) ψ_X is convex (that is, M_X is log-convex).

Proof. (a) To prove M_X is convex, we want to prove $\forall\alpha\in(0,1),a,b\in\mathrm{dom}(M_X)$, there is $M_X(\alpha a+(1-\alpha)b)\leq\alpha M_X(a)+(1-\alpha)M_X(b)$.

$$\begin{aligned}M_X(\alpha a+(1-\alpha)b)&=\mathbb{E}\left[\exp\left(\alpha a+(1-\alpha)bX\right)\right]\\ &\leq\mathbb{E}\left[\alpha\exp\left(aX\right)+(1-\alpha)\exp\left(bX\right)\right]\\ &=\alpha\mathbb{E}\left[\exp\left(aX\right)\right]+(1-\alpha)\mathbb{E}\left[\exp\left(bX\right)\right]=\alpha M_X(a)+(1-\alpha)M_X(b)\,,\end{aligned}$$

where the inequality comes from the convexity of $x\rightarrow\exp(x)$.

To prove $\mathrm{dom}(M_X)$ is an interval containing zero, we want to prove $M_X(0)<\infty$. It is obvious that $M_X(0)=\mathbb{E}\left[\exp(0\cdot X)\right]=\mathbb{E}\left[\exp(0)\right]=1<\infty$.

(b) For all $\lambda \in \mathrm{dom} M_X$, we have

$$M_X(\lambda) = \mathbb{E}\left[\exp(\lambda X)\right] \geq \exp\left(\mathbb{E}\left[\lambda X\right]\right) = \exp\left(\lambda \mathbb{E}\left[X\right]\right),$$

where the inequality comes from the convexity of $x \to \exp(x)$.

(c) TBD

(d) $M_X^{(1)}(\lambda) = \frac{d}{d\lambda}\mathbb{E}\left[\exp(\lambda X)\right] = \mathbb{E}\left[\frac{d}{d\lambda}\exp(\lambda X)\right] = \mathbb{E}\left[X\exp(\lambda X)\right]$. Recursively, we have $M_X^{(k)}(\lambda) = \mathbb{E}\left[X^k \exp(\lambda X)\right]$.

(e) According to the result of (d), we have $M_X^{(k)}(0) = \mathbb{E}\left[X^k \exp(0 \cdot X)\right] = \mathbb{E}\left[X^k\right]$.

(f) To prove ψ_X is convex, we want to prove $\forall \alpha \in (0,1), a, b \in \mathrm{dom}(\psi_X)$, there is $\psi_X(\alpha a + (1-\alpha)b) \leq \alpha\psi_X(a) + (1-\alpha)\psi_X(b)$.

$$\begin{aligned}
\psi_X(\alpha a + (1-\alpha)b) &= \log M_X(\alpha a + (1-\alpha)b) \\
&= \log \mathbb{E}\left[\exp\left(\alpha a + (1-\alpha)b\right)X\right] \\
&= \log \mathbb{E}\left[\exp\left(\alpha a X\right)\exp((1-\alpha)bX)\right] \\
&= \log \mathbb{E}\left[\left(\exp\left(aX\right)\right)^{\alpha}\left(\exp(bX)\right)^{(1-\alpha)}\right] \\
&\leq \log\left(\mathbb{E}\left[\exp\left(aX\right)\right]^{\alpha}\mathbb{E}\left[\exp(bX)\right]^{(1-\alpha)}\right) \\
&= \alpha \log \mathbb{E}\left[\exp\left(aX\right)\right] + (1-\alpha)\mathbb{E}\left[\exp(bX)\right] = \alpha\psi_X(a) + (1-\alpha)\psi_X(b),
\end{aligned}$$

where the inequality comes from the Hölder's inequality.

□

5.10(Large deviation theory) Let $X, X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed random variables with zero mean and moment-generating function M_X with $dom(M_X) = \mathbb{R}$. Let$\hat{\mu}_n = \frac{1}{n}\sum_{t=1}^{n} X_t$.

(a) Show that for any $\epsilon > 0$,

$$\frac{1}{n}\log \mathbb{P}\left(\hat{\mu}_n \geq \epsilon\right) \leq -\psi_X^*(\epsilon) = -\sup_{\lambda}(\lambda\epsilon - \log M_X(\lambda)) \tag{19}$$

(b) Show that when X is a Rademacher variable($\mathbb{P}\left(X = 1\right) = \mathbb{P}\left(X = -1\right) = \frac{1}{2}$),$\psi_X^*(\epsilon) = \frac{1+\epsilon}{2}\log(1+\epsilon) + \frac{1-\epsilon}{2}\log(1-\epsilon)$when$|\epsilon| \leq 1$ and $\psi_X^*(\epsilon) = +\infty$,otherwise.

(c) Show that when X is a centered Bernoulli random variable with parameter p (that is, $\mathbb{P}\left(X = -p\right) = 1 - p$ and $\mathbb{P}\left(X = 1-p\right) = p$) then $\psi_X^*(\epsilon) = \infty$ when ϵ is such that $p + \epsilon > 1$ and $\psi_X^*(\epsilon) = d(p+\epsilon, p)$ otherwise, where $d(p,q) = p\log(p/q) + (1-p)\log((1-p)/(1-q))$ is the relative entropy between the distributions $B(p)$ and $B(q)$.

(d) Show that when $X \sim N(0, \sigma^2)$ then $\psi_X^*(\epsilon) = \epsilon^2/(2\sigma^2)$.

(e) Let $\sigma^2 = \mathbb{V}[X]$. The (strong form of the) central limit theorem says that

$$\lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\hat{\mu}_n\sqrt{\frac{n}{\sigma^2}} \geq x\right) - (1 - \Phi(x))\right| = 0$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-y^2/2)dy$ is the cumulative distribution of the standard Gaussian. Let Z be a random variable distributed like a standard Gaussian. A careless application of this result might suggest that

$$\lim_{n\to\infty}\frac{1}{n}\log \mathbb{P}\left(\hat{\mu}_n \geq \epsilon\right) \stackrel{?}{=} \lim_{n\to\infty}\frac{1}{n}\log \mathbb{P}\left(Z \geq \epsilon\sqrt{\frac{n}{\sigma^2}}\right)$$

Evaluate the right-hand side. In light of the previous parts, what can you conclude about the validity of the question-marked equality? What goes wrong with the careless application of the central limit theorem? What do you conclude about the accuracy of this theorem?

Proof. (a) For all $\lambda \in dom M_X$,we have

$$\begin{aligned}
\frac{1}{n}\log \mathbb{P}\left(\hat{\mu}_n \geq \epsilon\right) &= \frac{1}{n}\log \mathbb{P}\left(\exp\left(\lambda n \hat{\mu}_n\right) \geq \exp\left(\lambda n \epsilon\right)\right) \\
&\leq \frac{1}{n}\log(\mathbb{E}\left[\exp\left(\lambda n \hat{\mu}_n\right)\right]\exp\left(-\lambda n \epsilon\right)) \\
&= \frac{1}{n}\log(\prod_{t=1}^{n}\mathbb{E}\left[\exp\left(\lambda X_t\right)\right]\exp\left(-\lambda n \epsilon\right)) \\
&= \frac{1}{n}\log(\mathbb{E}\left[\exp\left(\lambda X\right)\right]^n\exp\left(-\lambda n \epsilon\right)) \\
&= \log(\mathbb{E}\left[\exp\left(\lambda X\right)\right]\exp\left(-\lambda \epsilon\right)) \\
&= -(\lambda\epsilon - \log M_X(\lambda))
\end{aligned}$$

Since it holds for all $\lambda \in dom M_X$,we can imply

$$\frac{1}{n}\log \mathbb{P}\left(\hat{\mu}_n \geq \epsilon\right) \leq -\psi_X^*(\epsilon) = -\sup_{\lambda}(\lambda\epsilon - \log M_X(\lambda))$$

(b) By definition, $\psi_X(\lambda) = \log(\frac{1}{2}(\exp(-\lambda) + \exp(\lambda))) = \log(\cosh(\lambda))$, where $\cosh(\cdot)$ is the hyperbolic cosine function. To find the maximum of $f(\lambda) = \lambda\varepsilon - \psi_X(\lambda)$, we have $f'(\lambda) = \varepsilon - \tanh(\lambda)$, where $\tanh(\lambda) = \frac{\exp(\lambda)-\exp(-\lambda)}{\exp(\lambda)+\exp(-\lambda)}$ is the hyperbolic tangent function. Since $\tanh(\lambda) \in [-1, 1]$, $\sup_\lambda f(\lambda) = +\infty$ when $|\varepsilon| > 1$. Otherwise, we have

$$\begin{aligned}
\psi_X^*(\varepsilon) &= f\left(\tanh^{-1}(\varepsilon)\right) \\
&= \tanh^{-1}(\varepsilon)\varepsilon - \log\cosh\left(\tanh^{-1}(\varepsilon)\right) \\
&= \frac{\varepsilon}{2}\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right) + \frac{1}{2}\log\left(1-\varepsilon^2\right) \\
&= \frac{1+\varepsilon}{2}\log(1+\varepsilon) + \frac{1-\varepsilon}{2}\log(1-\varepsilon),
\end{aligned}$$

where the second equality holds as $\tanh^{-1}(\varepsilon) = \frac{1}{2}\log\left(\frac{1+\varepsilon}{1-\varepsilon}\right)$.

(c) By definition,

$$\begin{aligned}
\psi_X(\lambda) = \log\mathbb{E}\left[\exp(\lambda X)\right] &= \log\left(p\exp\left(\lambda(1-p)\right) + (1-p)\exp\left(-\lambda p\right)\right) \\
&= \log\left(p\exp\lambda\exp\left(-\lambda p\right) + (1-p)\exp\left(-\lambda p\right)\right) \\
&= \log\left(\exp\left(-\lambda p\right)\left(p\exp\lambda + 1 - p\right)\right) \\
&= -\lambda p + \log\left(p\exp\lambda + 1 - p\right).
\end{aligned}$$

Further,

$$\psi_X^*(\varepsilon) = \sup_{\lambda}\left(\lambda\varepsilon - \psi_X(\lambda)\right) = \sup_{\lambda}\left(\lambda\varepsilon + \lambda p - \log\left(p\exp\lambda + 1 - p\right)\right) \triangleq \sup_{\lambda} f(\lambda).$$

By letting $f'(\lambda) = \varepsilon + p - p\exp\lambda\frac{1}{p\exp\lambda+1-p} = 0$, we have $\lambda^* = \log\frac{(1-p)(p+\varepsilon)}{p(1-(p+\varepsilon))}$ when $p + \varepsilon < 1$.
Above all, when $p + \varepsilon < 1$, $\psi_X^*(\varepsilon) = f(\lambda^*) = d(p+\varepsilon, p)$.
When $p+\varepsilon = 1$, $\psi_X^*(\varepsilon) = \lim_{\lambda\to\infty} f(\lambda) = \lim_{\lambda\to\infty} \lambda - \log(p\exp\lambda + 1 - p) = \lim_{\lambda\to\infty} \lambda - \log(p\exp\lambda) = -\log p = d(p+\varepsilon, p)$.
When $p+\varepsilon > 1$, $\psi_X^*(\varepsilon) = \lim_{\lambda\to\infty} f(\lambda) = \lim_{\lambda\to\infty} \lambda(p+\varepsilon) - \log(p\exp\lambda + 1 - p) = \lim_{\lambda\to\infty} \lambda\varepsilon = \infty$.
The result has been proved.

□

5.11 (Hoeffding's lemma) Suppose that X is zero mean and $X \in [a, b]$ almost surely for constants $a < b$.

(a) Show that X is $(b-a)/2$-subgaussian.

(b) Prove Hoeffding's inequality (Lemma 2).

Lemma 2 (Hoeffding's inequality)**.** *For a zero-mean random variable X such that $X \in [a, b]$ almost surely for real values $a < b$, then $M_X(\lambda) \leq \exp(\lambda^2(b-a)^2/8)$. Applying the Cramér–Chernoff method shows that if $X_1, X_2, \ldots, X_n$ are independent and $X_t \in [a_t, b_t]$ almost surely with $a_t < b_t$ for all t. Then,*

$$\mathbb{P}\left(\frac{1}{n}\sum_{t=1}^{n}(X_t - \mathbb{E}[X_t]) \geq \varepsilon\right) \leq \exp\left(\frac{-2n^2\varepsilon^2}{\sum_{t=1}^{n}(b_t - a_t)^2}\right). \tag{20}$$

Proof. (a) To show X is $(b-a)/2$-subgaussian, we want to prove $\forall \lambda \in \mathbb{R}, \mathbb{E}[\exp(\lambda X)] \leq \exp\left(\frac{\lambda^2(b-a^2)}{8}\right)$.

According to the convexity of $x \to \exp(x)$ and Jensen's inequality, we have $\forall X \in [a, b]$,

$$\exp(\lambda X) = \exp\left(\frac{b-X}{b-a}\lambda a + \frac{X-a}{b-a}\lambda b\right) \leq \frac{b-X}{b-a}\exp(\lambda a) + \frac{X-a}{b-a}\exp(\lambda b).$$

Thus,

$$\begin{aligned}
\mathbb{E}[\exp(\lambda X)] &\leq \frac{b-\mathbb{E}[X]}{b-a}\exp(\lambda a) + \frac{\mathbb{E}[X]-a}{b-a}\exp(\lambda b) \\
&= \frac{b}{b-a}\exp(\lambda a) + \frac{-a}{b-a}\exp(\lambda b) \\
&= (1-\theta)\exp(\lambda a) + \theta\exp(\lambda b) \quad \left(\text{By letting } \theta = \frac{-a}{b-a}\right) \\
&= \exp(\lambda a)\left(1-\theta+\theta\exp(\lambda b - \lambda a)\right) \\
&= \exp(-\lambda(b-a)\theta)\left(1-\theta+\theta\exp(\lambda b - \lambda a)\right) \quad (\text{By representing } a \text{ using } \theta) \\
&= \exp\left(-\theta u + \log(1-\theta+\theta\exp(u))\right) \quad (\text{By letting } u = \lambda(b-a)) \\
&= \exp(\phi(u)), \quad \text{where} \quad \phi(u) = -\theta u + \log(1-\theta+\theta\exp(u)).
\end{aligned}$$

We next want to find the upper bound for $\exp(\phi(u))$. According to the Taylor's theorem with mean-values forms of the remainder, we have

$$\begin{aligned}
\phi(u) &= \phi(0) + u\phi'(0) + \frac{1}{2}u^2\phi''(v), \quad \text{where } v \in (0, u) \\
&= \frac{1}{2}u^2\phi''(v) \quad (\text{Since } \phi(0) = \phi'(0) = 0) \\
&= \frac{1}{2}u^2 \cdot \frac{\theta\exp(v)}{1-\theta+\theta\exp(v)}\left(1 - \frac{\theta\exp(v)}{1-\theta+\theta\exp(v)}\right) \\
&\leq \frac{1}{2}u^2 \cdot \frac{1}{4} = \frac{1}{8}u^2 = \frac{\lambda^2(b-a)^2}{8}.
\end{aligned}$$

Above all, we have proved that $\forall \lambda \in \mathbb{R}, \mathbb{E}[\exp(\lambda X)] \leq \exp(\phi(u)) \leq \exp\left(\frac{\lambda^2(b-a)^2}{8}\right)$.

(b) We only give the upper tail of the Hoeffding's Inequality using the Hoeffding's Lemma since the lower tail has a similar proof. Applying the Hoeffding' Lemma and the Chernoff bound technique immediately shows that

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\sum_{t=1}^{n}(X_t - \mathbb{E}[X_t]) \geq \epsilon\right) &= \mathbb{P}\left(\sum_{t=1}^{n}(X_t - \mathbb{E}[X_t]) \geq n\epsilon\right) \\
&\leq \mathbb{E}\left[\exp\left(\lambda\sum_{t=1}^{n}(X_t - \mathbb{E}[X_t])\right)\right]e^{-\lambda n\epsilon} \\
&= \left(\Pi_{t=1}^{n}\mathbb{E}[\exp\left(\lambda(X_t - \mathbb{E}[X_t])\right)]\right)e^{-\lambda n\epsilon} \\
&\leq \left(\Pi_{t=1}^{n}e^{\frac{\lambda^2(b_t-a_t)^2}{8}}\right)e^{-\lambda n\epsilon},
\end{aligned}$$

where $\lambda \geq 0$. Minimizing the RHS of the above inequality over λ shows that

$$\mathbb{P}\left(\frac{1}{n}\sum_{t=1}^{n}(X_t - \mathbb{E}[X_t]) \geq \epsilon\right) \leq \min_{\lambda \geq 0}\left(\Pi_{t=1}^{n} e^{\frac{\lambda^2(b_t-a_t)^2}{8}}\right) e^{-\lambda n \epsilon} = \exp\left(\frac{-2n^2\epsilon^2}{\sum_{t=1}^{n}(b_t-a_t)^2}\right).$$

□

5.14(Bernstein's inequality) Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables with $X_t - \mathbb{E}[X_t] \leq b$ almost surely and $S = \sum_{t=1}^{n} X_t - \mathbb{E}[X_t]$ and $v = \sum_{t=1}^{n} \mathbb{V}[X_t]$.

(a) Show that $g(x) = \frac{1}{2} + \frac{x}{3!} + \frac{x^2}{4!} + \cdots = \frac{\exp(x)-1-x}{x^2}$ is increasing.

(b) Let X be a random variable with $\mathbb{E}[X] = 0$ and $X \leq b$ almost surely. Show that $\mathbb{E}[\exp(X)] \leq 1 + g(b)\mathbb{V}[X]$.

(c) Prove that $(1+\alpha)\log(1+\alpha) - \alpha \geq \frac{3\alpha^2}{6+2\alpha}$ for all $\alpha \geq 0$. Prove that this is the best possible approximation in the sense that the 2 in the denominator cannot be increased.

(d) Let $\varepsilon > 0$ and $\alpha = b\varepsilon/v$ and prove that

$$\mathbb{P}(S \geq \varepsilon) \leq \exp\left(-\frac{v}{b^2}\left((1+\alpha)\log(1+\alpha) - \alpha\right)\right) \tag{21}$$

$$\leq \exp\left(-\frac{\varepsilon^2}{2v\left(1+\frac{b\varepsilon}{3v}\right)}\right). \tag{22}$$

(e) Use the previous result to show that

$$\mathbb{P}\left(S \geq \sqrt{2v\log\left(\frac{1}{\delta}\right)} + \frac{2b}{3}\log\left(\frac{1}{\delta}\right)\right) \leq \delta\,.$$

(f) Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables adapted to filtration $\mathbb{F} = (\mathcal{F}_t)_t$. Abbreviate $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \mid \mathcal{F}_t]$ and $\mu_t = \mathbb{E}_{t-1}[X_t]$. Define $S = \sum_{t=1}^{n} X_t - \mu_t$ and let $V = \sum_{t=1}^{n} \mathbb{E}_{t-1}[(X_t - \mu_t)^2]$ be the predictable variation of $\left(\sum_{t=1}^{p} X_t - \mu_t\right)_p$. Show that if $X_t - \mu_t \leq b$ holds almost surely for all $t \in [n]$, then with $\alpha = b\varepsilon/v$,

$$\mathbb{P}(S \geq \varepsilon, V \leq v) \leq \exp\left(-\frac{v}{b^2}\left((1+\alpha)\log(1+\alpha) - \alpha\right)\right).$$

Note that the right-hand side of this inequality is the same as that shown in Eq.(21).

Proof. (a) To show $g(x)$ is increasing, we want to show that $g'(x) \geq 0$.

$$g'(x) = \frac{(\exp(x)-1)x^2 - (\exp(x)-x-1)2x}{x^4} = \frac{x(x-2)\exp(x) + x(x+2)}{x^4}.$$

By letting $g'(x) = 0$, we find $x = 0$. And $\forall x \neq 0$, we have $g'(x) > 0$. Above all, we can show $g(x)$ is increasing.

(b) According to (a), we have $\exp(x) = x^2 g(x) + x + 1$. Thus

$$\begin{aligned}\mathbb{E}[\exp(X)] = \mathbb{E}\left[X^2 g(X) + X + 1\right] &= \mathbb{E}\left[X^2 g(X)\right] + \mathbb{E}[X] + 1 \\ &\leq g(b)\mathbb{E}\left[X^2\right] + \mathbb{E}[X] + 1 \\ &= g(b)\left(\mathbb{V}[X] + \mathbb{E}[X]^2\right) + \mathbb{E}[X] + 1 \\ &= g(b)\mathbb{V}[X] + 1\,,\end{aligned}$$

where the inequality is because $X \leq b$ and the increasing property of $g(x)$ proved in (a), the second last equality comes from the definition of $\mathbb{V}[X]$, the last equality holds since $\mathbb{E}[X] = 0$.

(c) Define $f(\alpha) = (1+\alpha)\log(1+\alpha) - \alpha - \frac{3\alpha^2}{6+2\alpha}$ and we want to prove $f(\alpha) \geq 0$ for any $\alpha \geq 0$.
We first compute $f'(\alpha)$ as

$$\begin{aligned} f'(\alpha) &= \log(1+\alpha) + (1+\alpha)\cdot\frac{1}{1+\alpha} - 1 - \frac{6\alpha^2+36\alpha}{(6+2\alpha)^2} \\ &= \log(1+\alpha) - \frac{3}{2} + \frac{54}{4(\alpha+3)^2}\,. \end{aligned}$$

It is obvious that $f'(0) = 0$. Further we find $f''(\alpha) = \frac{1}{1+\alpha} - \frac{27}{(\alpha+3)^3} \geq 0$ for any $\alpha \geq 0$. Thus we conclude that $f'(\alpha)$ increases in $[0,\infty)$ and $f'(\alpha) \geq 0$ in $[0,\infty)$. Above all, $f(\alpha)$ also increases in $[0,\infty)$ and $f(\alpha) \geq f(0) = 0, \forall\alpha \geq 0$.

(d)

$$\begin{aligned} \mathbb{P}\left(S \geq \varepsilon\right) &= \mathbb{P}\left(\exp(\lambda S) \geq \exp(\lambda\varepsilon)\right) \\ &\leq \exp(-\lambda\varepsilon)\mathbb{E}\left[\exp(\lambda S)\right] \\ &= \exp(-\lambda\varepsilon)\mathbb{E}\left[\exp\left(\lambda\sum_{t=1}^{n}(X_t - \mathbb{E}\left[X_t\right])\right)\right] \\ &= \exp(-\lambda\varepsilon)\prod_{t=1}^{n}\mathbb{E}\left[\exp(X_t - \mathbb{E}\left[X_t\right])\right] && (23) \\ &\leq \exp(-\lambda\varepsilon)\prod_{t=1}^{n}\left(1 + g(\lambda b)\lambda^2 V[X_t]\right) && (24) \\ &\leq \exp(-\lambda\varepsilon)\prod_{t=1}^{n}\exp\left(g(\lambda b)\lambda^2 V[X_t]\right) && (25) \\ &= \exp(-\lambda\varepsilon)\exp\left(g(\lambda b)\lambda^2\sum_{t=1}^{n}V[X_t]\right) \\ &= \exp(-\lambda\varepsilon)\exp\left(g(\lambda b)\lambda^2 v\right) \\ &= \exp(-\lambda\varepsilon)\exp\left(\frac{\exp(\lambda b) - 1 - \lambda b}{\lambda^2 b^2}\cdot\lambda^2 v\right) && (26) \\ &= \exp\left(\frac{v}{b^2}\left(\exp(\lambda b) - 1 - \lambda b\right) - \lambda\varepsilon\right) \\ &\leq \exp\left(-\frac{v}{b^2}\left((1+\alpha)\log(1+\alpha) - \alpha\right)\right) && (27) \\ &\leq \exp\left(-\frac{\varepsilon^2}{2v\left(1 + \frac{b\varepsilon}{3v}\right)}\right), && (28) \end{aligned}$$

where (23) is due to the independence of $(X_t)_{t\in[n]}$, (24) is because $X_t - \mathbb{E}\left[X_t\right] \leq b$ and the result of (b), (25) holds since $1 + x \leq \exp(x)$, (26) comes from the definition of $g(\lambda b)$. To get (27), we let $f(\lambda) = \frac{v}{b^2}\left(\exp(\lambda b) - 1 - \lambda b\right) - \lambda\varepsilon$ and find $f'(\lambda) = \frac{v}{b^2}\left(b\exp(\lambda b) - b\right) - \varepsilon$. By letting $f'(\lambda) = 0$, we have $f(\lambda) \leq f(\frac{1}{b}\log(1+\alpha)) = -\frac{v}{b^2}\left((1+\alpha)\log(1+\alpha) - \alpha\right)$. (28) holds according to (c).

(e) By choosing $\delta = \exp\left(-\frac{\varepsilon^2}{2v\left(1+\frac{b\varepsilon}{3v}\right)}\right)$, we have

$$\varepsilon^2 - 2v\log\left(\frac{1}{\delta}\right) - \frac{2b}{3}\varepsilon\log\left(\frac{1}{\delta}\right) = 0\,.$$

By solving this equality, we have

$$\varepsilon = \frac{1}{2}\sqrt{\left(\frac{2b}{3}\log\left(\frac{1}{\delta}\right)\right)^2 + 8v\log\left(\frac{1}{\delta}\right)} + \frac{b}{3}\log\left(\frac{1}{\delta}\right)$$

Then according to Eq.(22),

$$\begin{aligned}\mathbb{P}\left(S \geq \sqrt{2v\log\left(\frac{1}{\delta}\right)} + \frac{2b}{3}\log\left(\frac{1}{\delta}\right)\right) &= \mathbb{P}\left(S \geq \sqrt{2v\log\left(\frac{1}{\delta}\right)} + \frac{b}{3}\log\left(\frac{1}{\delta}\right) + \frac{b}{3}\log\left(\frac{1}{\delta}\right)\right) \\ &\leq \mathbb{P}\left(S \geq \frac{1}{2}\sqrt{\left(\frac{2b}{3}\log\left(\frac{1}{\delta}\right)\right)^2 + 8v\log\left(\frac{1}{\delta}\right)} + \frac{b}{3}\log\left(\frac{1}{\delta}\right)\right) \leq \delta\,,\end{aligned}$$

where the first inequality is because $\forall a, b \in \mathbb{R} : \sqrt{|a| + |b|} \leq \sqrt{|a|} + \sqrt{|b|}$

□

5.16 Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{P}(X_t \leq x) \leq x$ for each $x \in [0,1]$ and $t \in [n]$. Prove that for any $\varepsilon > 0$ that

$$\mathbb{P}\left(\sum_{t=1}^{n} \log(1/X_t) \geq \varepsilon\right) \leq \left(\frac{\varepsilon}{n}\right)^n \exp(n - \varepsilon).$$

Proof. In general we want to use the Cramer-Chernoff method:

$$\begin{aligned}\mathbb{P}\left(\sum_{t=1}^{n} \log(1/X_t) \geq \varepsilon\right) &= \mathbb{P}\left(\exp\left(\lambda \sum_{t=1}^{n} \log(1/X_i)\right) \geq \exp(\lambda\varepsilon)\right) \\ &\leq \exp(-\lambda\varepsilon)\mathbb{E}\left[\exp\left(\lambda \sum_{t=1}^{n} \log(1/X_t)\right)\right].\end{aligned}$$

To bound the expectation term, the assumption $\mathbb{P}(X_t \leq x) \leq x$ tells us that for each $\lambda > 0$,

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda \log(1/X_t)\right)\right] &= \int_0^\infty \mathbb{P}\left(\exp\left(\lambda \log(1/X_t)\right) \geq x\right) dx \\ &= 1 + \int_1^\infty \mathbb{P}\left(X_t \leq x^{-1/\lambda}\right) dx \leq 1 + \int_1^\infty x^{-1/\lambda} dx = \frac{1}{1-\lambda},\end{aligned}$$

which completes the proof when choosing $\lambda = (\varepsilon - n)/\varepsilon$. □

5.17 (Concentration for categorical distribution) Let $X_1, \ldots, X_n$ be an independent and identically distributed sequence taking values in $[m]$. For $i \in [m]$, let $p(i) = \mathbb{P}(X_1 = i)$ and $\hat{p}(i) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}\{X_t = i\}$. Show that for any $\delta \in (0,1)$,

$$\mathbb{P}\left(\|p - \hat{p}\|_1 \geq \sqrt{\frac{2[\log(\frac{1}{\delta}) + m\log(2)]}{n}}\right) \leq \delta$$

Hint Use the fact that $\|p - \hat{p}\|_1 = \max_{\lambda \in \{-1,1\}^m} \langle \lambda, p - \hat{p}\rangle$.

Proof. First, fix a $\lambda \in \{-1,1\}^m$ and use hint, then

$$\langle \lambda, p - \hat{p}\rangle = \frac{1}{n}\sum_{t=1}^{n} \langle p - e_{X_t}\rangle$$

where e_{X_t} means unit vector that is not 0 only at X_t.

$$\begin{aligned}|\langle \lambda, p - e_{X_t}\rangle| \leq \|\lambda\|_\infty \|p - e_{X_t}\|_1 \leq 2 \\ \mathbb{E}[\langle \lambda, p - e_{X_t}\rangle] = 0\end{aligned}$$

Then using Hoeffding's inequality:

$$\mathbb{P}(\langle \lambda, p - \hat{p}\rangle \geq \epsilon) \leq \exp\left(\frac{-2n^2\epsilon^2}{16n}\right) = \delta$$

which complete the proof when choosing $\epsilon = \sqrt{\frac{8\log(1/\delta)}{n}}$.

$$\mathbb{P}\left(\|p - \hat{p}\|_1 \geq \sqrt{\frac{8[\log(\frac{1}{\delta}) + m\log(2)]}{n}}\right) \leq \delta$$

$\square$

5.18 (Expectation of maximum) Let $X_1, \ldots, X_n$ be a sequence of σ-subgaussian random variables (possibly dependent) and $Z = \max_{t\in[n]} X_t$. Prove that

(a) $\mathbb{E}[Z] \leq \sqrt{2\sigma^2 \log(n)}$.

(b) $\mathbb{P}\left(Z \geq \sqrt{2\sigma^2 \log(n/\delta)}\right) \leq \delta$ for any $\delta \in (0,1)$.

Proof. (a) Let $\lambda > 0$. Then,

$$\exp(\lambda\mathbb{E}[Z]) \leq \mathbb{E}[\exp(\lambda Z)] \leq \sum_{t=1}^{n} \mathbb{E}[\exp(\lambda X_t)] \leq n\exp(\frac{\lambda^2\sigma^2}{2}).$$

Rearranging shows that

$$\mathbb{E}(Z) \leq \frac{log(n)}{\lambda} + \frac{\lambda\sigma^2}{2}.$$

Choosing $\lambda = \frac{1}{\sigma}\sqrt{2log(n)}$ shows that $\mathbb{E}(Z) \leq \sqrt{2\sigma^2 log(n)}$

(b) First notice that

$$\begin{aligned}\mathbb{P}\left(Z \geq \sqrt{2\sigma^2\log(n/\delta)}\right) &= \mathbb{P}\left(\exists i : X_i \geq \sqrt{2\sigma^2\log(n/\delta)}\right) \\ &\leq \sum_{i=1}^{n} \mathbb{P}\left(X_i \geq \sqrt{2\sigma^2\log(n/\delta)}\right),\end{aligned}$$

which is given directly by a union bound.

Then, according to Theorem 5.3, we have $\mathbb{P}\left(X_i \geq \sqrt{2\sigma^2\log(n/\delta)}\right) \leq \frac{\delta}{n}$ to complete the proof.

$\square$

Chapter 6 The Explore-Then-Commit Algorithm

6.1 (Subgaussian empirical estimates) Let π be the policy of ETC and $P_1, \dots, P_k$ be the 1-subgaussian distributions associated with the k arms. Provide a fully rigourous proof of the claim that

$$\hat{\mu}_i(mk) - \mu_i - \hat{\mu}_1(mk) + \mu_1$$

is $\sqrt{2/m}$-subgaussian. You should only use the definitions and the interaction protocol, which states that

1. $\mathbb{P}\left(A_t \in \cdot \mid A_1, X_1, \dots, A_{t-1}, X_{t-1}\right) = \pi\left(\cdot \mid A_1, X_1, \dots, A_{t-1}, X_{t-1}\right)$ a.s.
2. $\mathbb{P}\left(X_t \in \cdot \mid A_1, X_1, \dots, A_{t-1}, X_{t-1}, A_t\right) = P_{A_t}(\cdot)$ a.s.

Proof. By Lemma 5.4, it holds that $(\hat{\mu}_i(mk) - \mu_i)$ and $(\hat{\mu}_1(mk) - \mu_1)$ are both $\sqrt{1/m}$-subgaussian. Hence $(\hat{\mu}_i(mk) - \mu_i) - (\hat{\mu}_1(mk) - \mu_1)$ is $\sqrt{2/m}$-subgaussian, again according to Lemma 5.4 $\square$

6.2 (Minimax regret) Show that Eq. (6.6) implies the regret of an optimally tuned ETC for subgaussian two-armed bandits satisfies $R_n \leq \Delta + C\sqrt{n}$ where $C > 0$ is a universal constant.

Proof. We proceed by comparing the values of $n\Delta$ and $\Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}\right)$.

(a) If $n\Delta > \Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}\right)$, we have $(n-1)\Delta^2 > 4(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}) \geq 4$, which suggests that $\Delta \geq \frac{2}{\sqrt{n}}$. Therefore,

$$\begin{aligned}
R_n &= \Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}\right) \\
&= \Delta + \frac{4}{\Delta}\left(1 + \log\left(\frac{n\Delta^2}{4}\right)\right) \\
&= \Delta + \frac{4}{\Delta} + \frac{4}{\Delta}\log(\frac{n\Delta^2}{4}) \\
&\leq \Delta + 2\sqrt{n} + \frac{16}{e^4}\sqrt{n} \\
&= \Delta + C\sqrt{n},
\end{aligned}$$

where the inequality follows from taking $x^* = \frac{2e^4}{\sqrt{n}}$ to maximize $f(x) = \frac{4}{x}log(\frac{nx^2}{4})$.

(a) If $n\Delta \leq \Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}\right)$, we consider another two cases:

(i) If $\Delta \geq \frac{2}{\sqrt{n}}$, by (1) we still have $R_n = n\Delta \leq \Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{n\Delta^2}{4}\right)\right\}\right) \leq \Delta + C\sqrt{n}$.

(ii) If $\Delta < \frac{2}{\sqrt{n}}$, $R_n \leq n\Delta \leq 2\sqrt{n} \leq \Delta + C\sqrt{n}$, where the first inequality is trivial.

$\square$

6.3 Suppose $\Delta_1 = 0$, $\Delta_2 = \Delta > 0$. Then, the probability that we choose the suboptimal arm (i.e., the second arm) after commitment is

$$\begin{aligned}
\mathbb{P}(T_2(n) > m) &= \mathbb{P}(\hat{\mu}_2(2m) > \hat{\mu}_1(2m)) \\
&= \mathbb{P}([\hat{\mu}_2(2m) - \mu_2] - [\hat{\mu}_1(2m) - \mu_1] > \Delta \\
&\le \exp(-\frac{m\Delta^2}{4}),
\end{aligned}$$

where the inequality follows from Theorem 5.3. By letting $\exp(-\frac{m\Delta^2}{4}) = \delta$, we have $m = -\frac{4\log\delta}{\Delta^2}$. Hence, if we take $m = \min\{\lfloor\frac{n}{2}\rfloor, -\frac{4\log\delta}{\Delta^2}\}$, with high probability we have

$$\begin{aligned}
\bar{R}_n &= \Delta T_2(n) \\
&\le \Delta m \\
&= \min\{\lfloor\frac{n}{2}\rfloor\Delta, -\frac{4\log\delta}{\Delta}\}
\end{aligned}$$

6.4 (High-probability bounds (ii)) Repeat the previous exercise, but now prove a high probability bound on the random regret: $\hat{R}_n = n\mu^* - \sum_{t=1}^n X_t$. Compare this to the bound derived for the pseudo-regret in the previous exercise. What can you conclude?

Proof. Denote the reward received in the t-th interaction with arm i as $X_{i,t}$. From Ex. 6.3, we have that with probability $1-\delta$,

$$\begin{aligned}
\hat{R}_n &\le \sum_{t=1}^{n-m}(\mu_1 - X_{1,t}) + \sum_{t=1}^{m}(\mu_1 - X_{2,t}) \\
&= \sum_{t=1}^{n-m}(\mu_1 - X_{1,t}) + \sum_{t=1}^{m}(\mu_2 - X_{2,t}) + m\Delta\,.
\end{aligned}$$

Notice that the sum of the first two terms is ($\sqrt{(n-m)^2 + m^2}$)-subgaussian. Therefore, with probability $(1-\delta)^2$, we have $\hat{R}_n \le \sqrt{-2[(n-m)^2 + m^2]\log\delta} + m\Delta$. This suggests that compared to that derived for the pseudo-regret, the bound on the random regret is less tight with a smaller probability as more randomness is considered. □

6.5 Suppose $\Delta_1 = 0$, $\Delta_2 = \Delta > 0$.

(a) By Theorem 6.1, we have

$$\begin{aligned}
R_n(v) &= \Delta\mathcal{E}[T_2(n)] \\
&\le m\Delta + (n-2m)\Delta\exp(-\frac{m\Delta^2}{4}) \\
&\le m\Delta + n\Delta\exp(-\frac{m\Delta^2}{4}) \\
&\le m\Delta + n\sqrt{\frac{2}{m}}\exp(-\frac{1}{2}) \\
&= [\Delta + \sqrt{2}\exp(-\frac{1}{2})]n^{\frac{2}{3}},
\end{aligned}$$

where the last inequality follows from taking $x^* = \sqrt{\frac{2}{m}}$ to maximize $f(x) = x\exp(-\frac{m\Delta^2}{4})$, and the last equality follows from taking $m = n^{\frac{2}{3}}$.

Assume there is such a $C > 0$ that leads to $R_n(v) \le \Delta_v + Cn^{2/3}$ for any problem instance v and $n \ge 1$. Since trivially $R_n(v) \ge m\Delta$, we have $m\Delta \le \Delta + Cn^{\frac{2}{3}} \Rightarrow m \le 1 + \frac{Cn^{\frac{2}{3}}}{\Delta}$. Under this circumstance, we can easily find a problem instance v with $\Delta \to \infty$ such that $m \le 1$. Recalling we only explore $2m$ rounds, we will eventually pull the suboptimal arm with a high probability, which contradicts our assumption.

(c) We proceed by comparing the values of $\frac{C\log n}{\Delta}$ and $C\sqrt{n\log(n)}$.

(i) If $\Delta \geq \sqrt{\frac{\log n}{n}}$, $R_n(v) \leq \Delta + C\frac{\log n}{\Delta} \leq \Delta + C\sqrt{n\log n}$.

(ii) If $\Delta < \sqrt{\frac{\log n}{n}}$, $R_n(v) \leq n\Delta \leq \sqrt{n\log n} \leq \Delta + C\sqrt{n\log n}$.

(e) We proceed by comparing the values of e and $n\Delta^2$.

(i) If $\Delta \geq \sqrt{\frac{e}{n}}$, $R_n(v) \leq \Delta + \frac{C\log(n\Delta^2)}{\Delta} \leq \Delta + \frac{2C}{e}\sqrt{n} = \Delta + C\sqrt{n}$.

(ii) If $\Delta < \sqrt{\frac{e}{n}}$, $R_n(v) \leq n\Delta \leq \sqrt{en} \leq \Delta + C\sqrt{n}$.

6.6

The purpose of this exercise is to analyse a meta-algorithm based on the so-called **doubling trick** that converts a policy depending on the horizon to a policy with similar guarantees that does not. Let E be an arbitrary set of bandits. Suppose you are given a policy $\pi = \pi(n)$ designed for E that accepts the horizon n as a parameter and has a regret guarantee of

$$\max_{1\leq t\leq n} R_t(\pi(n), v) \leq f_n(v)\,,$$

where $f_n : \mathcal{E} \to [0, +\infty)$ is a sequence of function. Let $n_1 < n_2 < n_3 < \cdots$ be a fixed sequence of integers and consider the policy that runs π with horizon n_1 until round $t = \min\{n, n_1\}$, then runs π with horizon n_2 until $t = \min\{n, n_1 + n_2\}$, and then restarts again with horizon n_3 until $t = \min\{n, n_1 + n_2 + n_3\}$ and so-on. Note that t is the real-time counter and is not reset on each restart. Let π^* be the resulting policy. When $n_{l+1} = 2n_l$, the length of periods when π is used double with each phase, hence the name 'doubling trick'.

(a) Let $n > 0$ be arbitrary, $l_{\max} = \min\left\{l : \sum_{i=1}^{l} n_i \geq n\right\}$. Prove that for any $\nu \in \mathcal{E}$, the n-horizon regret of π^{dbl} on ν is at most

$$R_n(\pi^*, v) \leq \sum_{l=1}^{l_{\max}} f_{n_l}(v)$$

Proof.

$$R_n(\pi^*, v) = \sum_{l=1}^{l_{\max}} R_{n_l}(\pi(n_l), v) \leq \sum_{l=1}^{l_{\max}} \max_{1\leq t\leq n_l} R_t(\pi(n_l), v) \leq f_{n_l}(v) \leq \sum_{l=1}^{l_{\max}} f_{n_l}(v)$$

□

(b) Suppose that $f_n(\nu) \leq \sqrt{n}$. Show that if $n_l = 2^l - 1$, then for any $\nu \in \mathcal{E}$ and horizon n the regret of π^{dbl} is at most

$$R_n(\pi^*, v) \leq C\sqrt{n}\,,$$

where $C > 0$ is a carefully chosen universal constant.

Proof. Since $\sum_{i=1}^{l} n_i = \sum_{i=0}^{l-1} 2^i = 2^l - 1$, $l_{\max} = \lceil \log(n+1) \rceil$ and $2^{l_{\max}} \leq 2n+2$. Hence

$$R_n(\pi^*, v) \leq \sum_{l=1}^{l_{\max}} \sqrt{2}^{l-1} \leq \frac{1}{\sqrt{2}-1}\sqrt{2^{l_{\max}}} = 2(1+\sqrt{2})\sqrt{n}$$

□

(c) Suppose that $f_n(v) = g(v)\log(n)$ for some function $g : \mathcal{E} \to [0,\infty)$. What is the regret of π^* if $n_l = 2^l - 1$? Can you find a better choice of $(n_l)_l$?

$$R_n(\pi^*, v) \leq g(v)\sum_{l=1}^{l_{\max}} \log(2^{l-1}) = \frac{\log(2)}{2} g(v)(l_{\max}-1)l_{\max} \leq Cg(v)\log^2(n+1)$$

where C is some universal constant. A better choice is $n_l = 2^{2^{l-1}}$. With this,

$$R_n(\pi^*, v) \leq g(v)\sum_{l=1}^{l_{\max}} \log(2^{2^{l-1}}) \leq \log(2)g(v)2^{l_{\max}} \leq Cg(v)\log(n)$$

where $C > 0$ is another universal constant.

6.7 (ε-GREEDY) For this exercise assume the rewards are 1 -subgaussian and there are $k \geq 2$ arms. The ε-greedy algorithm depends on a sequence of parameters $\varepsilon_1, \varepsilon_2, \ldots$. First it chooses each arm once and subsequently chooses $A_t = \operatorname{argmax}_i \hat{\mu}_i(t-1)$ with probability $1-\varepsilon_t$ and otherwise chooses an arm uniformly at random.

(a) Prove that if $\varepsilon_t = \varepsilon > 0$, then $\lim_{n\to\infty} \frac{R_n}{n} = \frac{\varepsilon}{k}\sum_{i=1}^k \Delta_i$.

(b) Let $\Delta_{\min} = \min\{\Delta_i : \Delta_i > 0\}$ and let $\varepsilon_t = \min\left\{1, \frac{Ck}{t\Delta_{\min}^2}\right\}$, where $C > 0$ is a sufficiently large universal constant. Prove that there exists a universal $C' > 0$ such that

$$R_n \leq C' \sum_{i=1}^k \left(\Delta_i + \frac{\Delta_i}{\Delta_{\min}^2} \log\max\left\{e, \frac{n\Delta_{\min}^2}{k}\right\}\right)$$

Proof. (a) To prove the $\frac{R_n}{n}$ in ε-GREEDY algorithms, we prove it with uniformly choosing arm and select $A_t = \operatorname{argmax}_i \hat{\mu}_i(t-1)$.

1 In uniformly choosing arm situation (with ε probability), the regret is

$$\begin{aligned}
&\sum_{t=1}^n (x - x_{at}) \\
&= \sum_{i=1}^n \mathbb{E}(x - x_{at}) \\
&= \sum_{i=1}^n \mathbb{E}\Delta_i \\
&= \sum_{t=1}^n \frac{\sum_{i=1}^k \Delta_i}{k}.
\end{aligned}$$

$\lim_{n\to\infty} \frac{R_n}{n} = \frac{\sum_{i=1}^k \Delta_i}{k}$

2 In selecting $A_t = \operatorname{argmax}_i \hat{\mu}_i(t-1)$ situation, assuming the best arm is arm 1.
We could prove the select arm is always the best arm 1, when t is large enough. Namely,
Prove $\exists t_n, t > t_n, A_t = 1$
By contradiction, assuming for $\forall t_n, \exists A_t \neq 1$ in $t > t_n$.
Since $A_t \neq 1$ and $A_t = \operatorname{argmax}_i \hat{\mu}_i(t-1)$, $\hat{\mu}_i > \hat{\mu}_1$.
According to reward is 1-subgaussian,

$$|\hat{\mu}_i - \mu_i| < \delta_i$$

$$|\hat{\mu}_1 - \mu_1| < \delta_1$$

Thus

$$\hat{\mu}_i < \mu_i + \delta_i$$

$$\hat{\mu}_1 > \mu_1 - \delta_1$$

Let t_n big enough to enable $\delta_i + \delta_1 < \mu_i - \mu_1$.
$\mu_i + \delta_i < \mu_1 - \delta_1$
$\hat{\mu}_i < \hat{\mu}_1$ and $A_t = 1$ which is contradict with our assumption. Thus $\exists t_n, t > t_n, A_t = 1$.
When $t > t_n, A_1 = 1$, $R_t - R_{t_n} = 0$;
$R_n = R_{tn} + R_t - R_n \leq t_n \Delta$

$$\lim_{n\to\infty} \frac{R_n}{n} = 0$$

The algorithm uniformly chooses arm with ε probability and selects $A_t = \text{argmax}_i\, \hat{\mu}_i(t-1)$ with $1-\varepsilon$ probability.
$\lim_{n\to\infty} \frac{R_n}{n} = \varepsilon \frac{\sum_{i=1}^k \Delta_i}{k} + (1-\varepsilon)\cdot 0 = \frac{\varepsilon}{k}\sum_{i=1}^k \Delta_i$

(b)

□

6.8 (Elimination algorithm) A simple way to generalise the ETC policy to multiple arms and overcome the problem of tuning the commitment time is to use an elimination algorithm. The algorithm operates in phases and maintains an active set of arms that could be optimal. In the ℓ-th phase, the algorithm aims to eliminate from the active set all arms i for which $\Delta_i \geq 2^{-\ell}$.

Without loss of generality, assume that arm 1 is an optimal arm. You may assume that the horizon n is known.

(a) Show that for any $\ell \geq 1$,

$$\mathbb{P}(1 \notin A_{\ell+1}, 1 \in A_\ell) \leq k \exp\left(-\frac{m_\ell 2^{-2\ell}}{4}\right).$$

(b) Show that if $i \in [k]$ and $\ell \geq 1$ are such that $\Delta_i \geq 2^{-\ell}$, then

$$\mathbb{P}(i \in A_{\ell+1}, 1 \in A_\ell, i \in A_\ell) \leq \exp\left(-\frac{m_\ell\left(\Delta_i - 2^{-\ell}\right)^2}{4}\right).$$

(c) Let $\ell_i = \min\left\{\ell \geq 1 : 2^{-\ell} \leq \Delta_i/2\right\}$. Choose m_ℓ in such a way that $\mathbb{P}(\text{ exists } \ell : 1 \notin A_\ell) \leq 1/n$ and $\mathbb{P}(i \in A_{\ell_i+1}) \leq 1/n$.

(d) Show that your algorithm has regret at most

$$R_n \leq C \sum_{i:\Delta_i>0} \left(\Delta_i + \frac{1}{\Delta_i}\log(n)\right),$$

where $C > 0$ is a carefully chosen universal constant.

Proof. (a) By applying concentration for subgaussian random variables, we have:

$$\begin{aligned}\mathbb{P}(1 \notin A_{\ell+1}, 1 \in A_\ell) &\leq \mathbb{P}\left(1 \in A_\ell,\ \text{exists } i \in A_\ell\backslash\{1\} : \hat{\mu}_{i,\ell} \geq \hat{\mu}_{1,\ell} + 2^{-\ell}\right) \\ &= \mathbb{P}\left(1 \in A_\ell,\ \text{exists } i \in A_\ell\backslash\{1\} : \hat{\mu}_{i,\ell} - \hat{\mu}_{1,\ell} \geq 2^{-\ell}\right) \\ &\leq k \exp\left(-\frac{m_\ell 2^{-2\ell}}{4}\right),\end{aligned}$$

where in the last final inequality we used (c) of Lemma 5.4 and Theorem 5.3.

(b) Again, concentration gives that:

$$\begin{aligned}\mathbb{P}\left(i \in A_{\ell+1}, 1 \in A_\ell, i \in A_\ell\right) &\leq \mathbb{P}\left(1 \in A_\ell, i \in A_\ell, \hat{\mu}_{i,\ell} + 2^{-\ell} \geq \hat{\mu}_{1,\ell}\right) \\ &= \mathbb{P}\left(1 \in A_\ell, i \in A_\ell, (\hat{\mu}_{i,\ell} - \mu_i) - (\hat{\mu}_{1,\ell} - \mu_1) \geq \Delta_i - 2^{-\ell}\right) \\ &\leq \exp\left(-\frac{m_\ell \left(\Delta_i - 2^{-\ell}\right)^2}{4}\right).\end{aligned}$$

(c) We first define

$$m_\ell = 2^{4+2\ell} \log(\ell/\delta),$$

where $\delta \in (0,1)$ is some constant to be chosen late. Then by Part (a),

$$\begin{aligned}\mathbb{P}\left(\text{ exists } \ell : 1 \notin A_\ell\right) &\leq \sum_{\ell=1}^{\infty} \mathbb{P}\left(1 \notin A_{\ell+1}, 1 \in A_\ell\right) \\ &\leq k \sum_{\ell=1}^{\infty} \exp\left(-\frac{m_\ell 2^{2\ell}}{4}\right) \\ &\leq k\delta \sum_{\ell=1}^{\infty} \frac{1}{\ell^2} \\ &= \frac{k\pi^2\delta}{6}.\end{aligned}$$

Furthermore, by Part (b),

$$\begin{aligned}\mathbb{P}\left(i \in A_{\ell_i+1}\right) &\leq \mathbb{P}\left(i \in A_{\ell_i+1}, i \in A_{\ell_i}, 1 \in A_{\ell_i}\right) + \mathbb{P}\left(1 \notin A_{\ell_i}\right) \\ &\leq \exp\left(-\frac{m_\ell \left(\Delta_i - 2^{-\ell_i}\right)^2}{4}\right) + \frac{k\pi^2\delta}{6} \\ &\leq \exp\left(-\frac{m_\ell 2^{-2\ell_i}}{16}\right) + \frac{k\pi^2\delta}{6} \\ &\leq \delta\left(1 + \frac{k\pi^2}{6}\right).\end{aligned}$$

Choosing $\delta = n^{-1}\left(1 + k\pi^2/6\right)^{-1}$ completes the result.

(d) For $n < k$, all actions need to be tried at most once leading to a trivial result. For $n \geq k$, let i be a suboptimal action. Notice that $2^{-\ell_i} \geq \Delta_i/4$ and hence $2^{2\ell_i} \leq 16/\Delta_i^2$. Furthermore, $m_\ell \geq m_1 \geq 1$ for $\ell \geq 1$. Therefore, we have

$$\begin{aligned}\mathbb{E}\left[T_i(n)\right] &\leq n\mathbb{P}\left(i \in A_{\ell_i+1}\right) + \sum_{\ell=1}^{\ell_i \wedge n} m_\ell \\ &\leq 1 + \sum_{\ell=1}^{\ell_i \wedge n} 2^{4+2\ell} \log\left(\frac{n}{\delta}\right) \\ &\leq 1 + C2^{2\ell_i} \log(nk) \\ &\leq 1 + \frac{16C}{\Delta_i^2} \log(nk)\end{aligned}$$

where $C > 1$ is a suitably large universal constant derived by naively bounding the logarithmic term and the geometric series. The result follows from upper bounding $\log(nk) \leq 2\log(n)$ which follows from $k \leq n$ and the standard regret decomposition, as described in Lemma 4.5.

$\square$

Chapter 7 The Upper Confidence Bound Algorithm

7.1 (Concentration for sequences of random length) In this exercise, we investigate one of the more annoying challenges when analyzing sequential algorithms. Let $X_1, X_2, \dots$ be a sequence of independent standard Gaussian random variables defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $T : \Omega \to \{1, 2, 3, \dots\}$ is another random variable, and let $\hat{\mu} = \sum_{t=1}^{T} X_t / T$ be the empirical mean based on T samples.
7.2 (Relaxing the subgaussian assumption) In this chapter, we assumed the pay-off distributions were 1-subgaussian. The purpose of this exercise is to relax this assumption.

(a) First suppose that σ^2 is a known constant and that $\nu \in \mathcal{E}_{SG}^k(\sigma^2)$. Modify the UCB algorithm and state and prove an analogue of Theorems 7.1 and 7.2 for this case.

(b) Now suppose that $\nu = (P_i)_{i=1}^k$ is chosen so that P_i is σ_i-subgaussian where $(\sigma_i^2)_{i=1}^k$ are known. Modify the UCB algorithm and state and prove an analogue of Theorems 7.1 and 7.2 for this case.

(c) If you did things correctly, the regret bound in the previous part should not depend on the values $\left\{\sigma_i^2 : \Delta_i = 0\right\}$. Explain why not.

Proof. (a) We modify the UCB algorithm by replacing $\sqrt{\frac{2\log(1/\delta)}{T_i(t-1)}}$ with $\sqrt{\frac{2\sigma^2\log(1/\delta)}{T_i(t-1)}}$ in Eq. (7.2) on page 85 of the book. Recall the reward $X_{t,i}$ of each arm i at time t is σ-subgaussian. According to the properties of subgaussian random variables, we have $\mathbb{P}\left(|\hat{\mu}_{is} - \mu_i| \geq \sqrt{\frac{2\sigma^2\log(1/\delta)}{s}}\right) \leq 2\delta$ for each arm i and $s > 0$.

Define event $\mathcal{F}_t = \left\{\exists i \in [k] : |\hat{\mu}_i(t) - \mu_i| \geq \sqrt{\frac{2\sigma^2\log(1/\delta)}{T_i(t-1)}}\right\}$. The regret can be decomposed by

$$
\begin{aligned}
R_n &= \sum_{i:\Delta_i>0} \Delta_i \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{A_t = i\}\right] \\
&\leq \sum_{i:\Delta_i>0} \Delta_i \left\{\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{A_t = i, \mathcal{F}_t\}\right] + \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{A_t = i, \neg\mathcal{F}_t\}\right]\right\} \\
&\leq \sum_{i:\Delta_i>0} \Delta_i \left\{\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{A_t = i, \neg\mathcal{F}_t\}\right] + 2nk\delta\right\} \\
&\leq \sum_{i:\Delta_i>0} \Delta_i \left\{\mathbb{E}\left[1 + \sum_{s=1}^{n-1} \mathbb{1}\left\{\mu_i + 2\sqrt{\frac{2\sigma^2\log(1/\delta)}{s}}\right\} \geq \mu_1\right] + 2nk\delta\right\} \\
&= \sum_{i:\Delta_i>0} \Delta_i \left(3 + \frac{16\sigma^2\log n}{\Delta_i^2}\right),
\end{aligned}
$$

where the last inequality holds by choosing $\delta = 1/n^2$. Thus the analogue of Theorems 7.1 for this case has been proved. Similar to the proof technique of Theorem 7.2, by setting $\Delta = 4\sqrt{\frac{k\sigma^2\log n}{n}}$, we have

$R_n \le 8\sqrt{nk\sigma^2 \log n} + 3\sum_{i:\Delta_i>0} \Delta_i$. Thus the analogue of Theorems 7.2 for this case has also been proved.

(b) For this case, we modify the UCB algorithm by replacing $\sqrt{\frac{2\log(1/\delta)}{T_i(t-1)}}$ with $\sqrt{\frac{2\sigma_i^2 \log(1/\delta)}{T_i(t-1)}}$ in Eq. (7.2) on page 85 of the book. The same proof techniques with part (a) can be used to derive the regret upper bound of order $R_n = \sum_{i:\Delta_i>0} \Delta_i \left(3 + \frac{16\sigma_i^2 \log n}{\Delta_i^2}\right)$, which is the analogue of Theorems 7.1 for this case. Define $\sigma_{\max} = \max_{i:\Delta_i>0} \sigma_i$ and set $\Delta = 4\sqrt{\frac{k\sigma_{\max}^2 \log n}{n}}$, we can get the analogue of Theorems 7.2 for this case with $R_n \le 8\sqrt{nk\sigma_{\max}^2 \log n} + 3\sum_{i:\Delta_i>0} \Delta_i$.

(c) As shown in part (a) and (b), our regret bound does not depend on the values $\left\{\sigma_i^2 : \Delta_i = 0\right\}$. Due to the monotonicity between UCB_1 and μ_1, the number of pulls of a suboptimal arm i only depends on σ_i but is not influenced by optimal arms.

□

7.4 (Phased UCB (i)) Fix a 1-subgaussian k-armed bandit environment and a horizon n. Consider the version of UCB that works in phases of exponentially increasing length of $1, 2, 4, \ldots$. In each phase, the algorithm uses the action that would have been chosen by UCB at the beginning of the phase (see Algorithm 4).

(a) State and prove a bound on the regret for this version of UCB.

(b) Compare your result with Theorem 7.1.

(c) How would the result change if the ℓ-th phase had a length of $\lceil \alpha^\ell \rceil$ with $\alpha > 1$?

Proof. (a) Let $T_i(n)$ denote the number of times pulling arm i. By definition we have

$$\begin{aligned}
\mathbb{E}\left[T_i(n)\right] &= \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\left\{A_t = i\right\}\right] \\
&= \mathbb{E}\left[\sum_{\ell=1}^{\ell_{max}} 2^\ell \mathbb{1}\left\{A_\ell = i\right\}\right] \\
&= \mathbb{E}\left[\sum_{\ell=1}^{\ell_{max}} 2^\ell \mathbb{1}\left\{A_\ell = i, \text{UCB}_i > \mu_1\right\}\right] \\
&= \mathbb{E}\left[\sum_{\ell=1}^{\ell_{max}} 2^\ell \mathbb{1}\left\{A_\ell = i, \mu_i + \sqrt{\frac{2\log(1/\delta)}{T_i(t-1)}} > \mu_1\right\}\right] \\
&= \mathbb{E}\left[\sum_{\ell=1}^{\ell_{max}} 2^\ell \mathbb{1}\left\{A_\ell = i, T_i(t-1) < \frac{8\log\frac{1}{\delta}}{\Delta_i^2}\right\}\right],
\end{aligned}$$

where Δ_i is the suboptimality gap.

To find ℓ_i such that $\sum_{\ell=1}^{\ell_i} 2^\ell \ge \frac{8\log\frac{1}{\delta}}{\Delta_i^2}$, we derive that $\ell_i \ge \lceil \log_2 \frac{4\log\frac{1}{\delta}}{\Delta_1^2} + 1 \rceil$. Hence, we have

$$\mathbb{E}\left[T_i(n)\right] \le \sum_{\ell=1}^{\ell_i} 2^\ell \le 32\frac{\log\frac{1}{\delta}}{\Delta_i^2},$$

Taking $\delta = \frac{1}{n^2}$ results in an upper bound $64\frac{\log n}{\Delta_i^2}$ for $\mathbb{E}\left[T_i(n)\right]$.

(b) Compared with Theorem 7.1, the constant is larger while the order remains the same.

(c) With the alternative phase length, the above analysis results in an upper bound of $\mathbb{E}\left[T_i(n)\right]$:

$$\mathbb{E}\left[T_i(n)\right] \leq \sum_{\ell=1}^{\ell_i} \alpha^\ell \leq 16\frac{\alpha^2 \log n}{\Delta_i^2},$$

where $\ell_i = \left[\log_\alpha \frac{8(\alpha-1)\log\frac{1}{\delta}}{\alpha\Delta_i^2} + 1\right]$.

□

7.6

$$\begin{aligned}
\hat{\delta}^2 &= \frac{1}{n}\sum_{t=1}^{n}(\hat{\mu} - X_t)^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}[(\hat{\mu} - \mu) + (\mu - X_t)]^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}(\hat{\mu} - \mu)^2 + \frac{1}{n}\sum_{t=1}^{n}(\mu - X - t)^2 + \frac{2}{n}\sum_{t=1}^{n}(\hat{\mu} - \mu)(\mu - X_t)
\end{aligned}$$

$\because \hat{\mu} - \mu = \frac{1}{n}\sum_{t=1}^{n} X_t - \mu = \frac{1}{n}\sum_{t=1}^{n}(X_t - \mu)$

$$\begin{aligned}
\therefore \frac{2}{n}\sum_{t=1}^{n}(\hat{\mu} - \mu)(\mu - X - t) &= \frac{2}{n}\sum_{t=1}^{n}[\frac{1}{n}\sum_{t=1}^{n}(X_t - \mu)](\mu - X_t) \\
&= -2[\frac{1}{n}\sum_{t=1}^{n}(X_t - \mu)]^2 \\
&= -2[\hat{\mu} - \mu]^2
\end{aligned}$$

$$\begin{aligned}
\hat{\delta}^2 &= (\hat{\mu} - \mu)^2 + \frac{1}{n}\sum_{t=1}^{n}(\mu - X_t)^2 - 2(\hat{\mu} - \mu)^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}(\mu - X_t)^2 - (\hat{\mu} - \mu)^2
\end{aligned}$$

Chapter 8

8.1

$$
\begin{aligned}
\sum_{s=1}^{n} exp(-\frac{s\varepsilon^2}{2}) &= exp(-\frac{\varepsilon^2}{2}) + ... + exp(-\frac{n\varepsilon^2}{2}) \\
&= \frac{exp(-\frac{\varepsilon^2}{2})[exp(-\frac{n\varepsilon^2}{2}) - 1]}{exp(-\frac{\varepsilon^2}{2}) - 1} \\
&= \frac{exp(-\frac{\varepsilon^2}{2})[1 - exp(-\frac{n\varepsilon^2}{2})]}{1 - exp(-\frac{\varepsilon^2}{2})} \\
&\leq \frac{exp(-\frac{\varepsilon^2}{2})}{1 - exp(-\frac{\varepsilon^2}{2})} \\
&\leq \frac{2}{\varepsilon^2}
\end{aligned}
$$

$\because f(t) = 1 + tlog^2(t) \therefore \frac{1}{f(t)} = \frac{1}{1+tlog^2(t)} \leq \frac{1}{tlog^2(t)}$

$$
\begin{aligned}
\sum_{t=1}^{n} \frac{1}{f(t)} &\leq \sum_{t=1}^{20} + \int_{20}^{\infty} \frac{dt}{f(t)} \\
&\leq \sum_{t=1}^{20} + \int_{20}^{\infty} \frac{dt}{tlog(t)} \\
&= \sum_{t=1}^{20} + \frac{1}{log(20)} \\
&\leq \frac{5}{2}
\end{aligned}
$$

8.2

$$
\begin{aligned}
&\mathbb{E}\left[T_2(n)\right] \\
&= \mathbb{E}\left[\mathbb{1}\{A_t = 2\}\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\left\{\hat{\mu}_1(t-1) + \sqrt{\frac{2\log(f(t))}{T_1(t-1)}} \leq \mu_1 - \epsilon\right\}\right] + \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\left\{\hat{\mu}_2(t-1) + \sqrt{\frac{2\log(f(t))}{T_2(t-1)}} \leq \mu_1 - \epsilon\right\}\right]
\end{aligned}
$$

Since $\mu_2 = 0$ is known, we can rewrite that

$$
\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mu_2 \geq \mu_1 - \epsilon\}\right] = \mathbb{E}\left[0 \geq \mu_1 - \epsilon\right].
$$

Since $\mu_1 > 0$, when $\epsilon \to 0$, $\mathbb{E}\left[0 \geq \mu_1 - \epsilon\right] = 0$. We take $\epsilon = \log^{-\frac{1}{4}}(n)$.

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\left\{\hat{\mu}_1(t-1) + \sqrt{\frac{2\log(f(t))}{T_1(t-1)}} \leq \mu_1 - \epsilon\right\}\right] \\
\leq &\sum_{t=1}^{n}\sum_{s=1}^{n} \mathbb{P}\left(\hat{\mu}_{1,s} + \sqrt{\frac{2\log(f(t))}{s}} \leq \mu_1 - \epsilon\right) \\
\leq &\sum_{t=1}^{n}\sum_{s=1}^{n} \exp(-\frac{\delta(\sqrt{\frac{2\log(f(t))}{s}} + \epsilon)^2}{2}) \\
\leq &\sum_{t=1}^{n} \frac{1}{f(t)} \sum_{s=1}^{n} \exp(-\frac{s\epsilon^2}{2}) \\
\leq &\frac{5}{\epsilon^2}
\end{aligned}
$$

Then we have $R_n = \mu_1 \mathbb{E}\left[T_2(n)\right] \leq \frac{5}{\epsilon^2} = 5\log^{\frac{1}{2}} n$, and $\lim_{n\to\infty} \sup \frac{R_n}{\log n}$=0.

Chapter 9 The Upper Confidence Bound Algorithm: Minimax Optimality

9.1 (Submartingale property) Let $X_1, X_2, \ldots, X_n$ be adapted to filtration $\mathbb{F} = (\mathcal{F}_t)_t$ with $\mathbb{E}[X_t \mid \mathcal{F}_{t-1}] = 0$ almost surely. Prove that $M_t = \exp\left(\lambda \sum_{s=1}^t X_s\right)$ is a $\mathbb{F}$-submartingale for any $\lambda \in \mathbb{R}$.

Proof. Notice that $M_t = \exp(\lambda \sum_{s=1}^t X_s) = \exp(\lambda \sum_{s=1}^{t-1} X_s) \exp(\lambda X_t) = \exp(\lambda X_t) M_{t-1}$. Therefore we have

$$\begin{aligned}
\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] &= \mathbb{E}[\exp(\lambda X_t) M_{t-1} \mid \mathcal{F}_{t-1}] \\
&= M_{t-1} \mathbb{E}[\exp(\lambda X_t) \mid \mathcal{F}_{t-1}] \\
&\geq M_{t-1} \exp(\lambda \mathbb{E}[X_t \mid \mathcal{F}_{t-1}]) \\
&= M_{t-1},
\end{aligned}$$

where the inequality follows from Jenson's inequality. □

9.2 Let $\Delta_{\min} = \min_{i:\Delta_i > 0} \Delta_i$. Show there exists a universal constant $C > 0$ such that the regret of MOSS is bounded by

$$R_n \leq Ck\Delta_{\min} \log^+(\frac{n\Delta_{\min}^2}{k}) + \sum_{i=1}^k \Delta_i$$

Proof.

$$\begin{aligned}
R_n &= \sum_{i=1}^k \Delta_i \mathbb{E}[T_i(n)] \\
&\leq \Delta_i + \Delta_i \mathbb{E}\left[\sum_{i=k+1}^n \mathbb{1}\left\{A_t = i; UCB_i(t) > \mu_i + \frac{\Delta_i}{2}\right\}\right] + \Delta_i \mathbb{E}\left[\sum_{i=k+1}^n \mathbb{1}\left\{UCB_1(t) \leq \mu_1 - \frac{\Delta_i}{2}\right\}\right] \\
&\leq \Delta_i + \underbrace{\Delta_i \mathbb{E}\left[\sum_{s=1}^n \mathbb{1}\left\{\hat{\mu}_{i,s} + \sqrt{\frac{4}{s}\log^+(\frac{n}{ks})} > \mu_i + \frac{\Delta}{2}\right\}\right]}_{(a)} + \underbrace{\Delta_i \mathbb{E}\left[\sum_{i=k+1}^n \mathbb{1}\left\{UCB_1(t) \leq \mu_1 - \frac{\Delta_i}{2}\right\}\right]}_{(b)}
\end{aligned}$$

To bound (a), from lemma 8.2 we have that

$$\begin{aligned}
&\Delta_i \mathbb{E}\left[\sum_{s=1}^n \mathbb{1}\left\{\hat{\mu}_{i,s} + \sqrt{\frac{4}{s}\log^+(\frac{n}{ks})} > \mu_i + \frac{\Delta}{2}\right\}\right] \\
&\leq \Delta_i (1 + \frac{4}{\Delta_i^2}(\log^+ \frac{n\Delta_i^2}{K} + \sqrt{\pi \log^+ \frac{n\Delta_i^2}{K}} + 1)) \\
&= O(\frac{1}{\Delta_i} \log^+ \frac{n\Delta_i^2}{K})
\end{aligned}$$

Then we turn to bound (b), using lemma 9.3 we have that

$$\begin{aligned}
&\Delta_i \mathbb{E}\left[\sum_{i=k+1}^{n} \mathbb{1}\left\{UCB_1(t) \leq \mu_1 - \frac{\Delta_i}{2}\right\}\right] \\
&= \Delta_i \sum_{t=k+1}^{n} \mathbb{P}\left(UCB_1(t) \leq \mu_1 - \frac{\Delta_i}{2}\right) \\
&\leq \frac{60K}{\Delta_i} = O(\frac{K}{\Delta_i})
\end{aligned}$$

Thus we can have

$$\Delta_i \mathbb{E}\left[T_i(n)\right] \leq O(\Delta_i + \frac{\log^+ \frac{n\Delta_i^2}{K}}{\Delta_i} + \frac{K}{\Delta_i}) = O(\Delta_i + \frac{\log^+ \frac{n\Delta_i^2}{K}}{\Delta_i})$$

$$R_n = \sum_{i=1}^{k} \Delta_i \mathbb{E}\left[T_i(n)\right] \leq O(\sum_{i=1}^{k} \Delta_i + \frac{K \log^+ \frac{n\Delta_{\min}^2}{K}}{\Delta_{\min}})$$

$\square$

Chapter 10 The Upper Confidence Bound Algorithm: Bernoulli Noise

10.1 (Pinsker's inequality) Prove Lemma 10.2(b).
Hint Consider the function $g(x) = d(p, p + x) - 2x^2$ over the $[-p, 1-p]$ interval. By taking derivatives, show that $g \geq 0$.

Proof. As the hint suggested, we validate the non-negativity of $g(x)$ over $[-p, 1-p]$.

$$\begin{aligned} g'(x) &= -\frac{p}{p+x} + \frac{1-p}{1-p-x} - 4x \\ &= \frac{x}{(1-p-x)(p+x)} - 4x \\ &\geq 0, \end{aligned}$$

which leads to $g(q-p) = d(p,q) - 2(p-q)^2 \geq 0$ and finally $d(p,q) \geq 2(p-q)^2$. □

10.2 (Asymptotic optimality) Prove the asymptotic claim in Theorem 10.6.
Hint Choose $\varepsilon_1, \varepsilon_1$ to decrease slowly with n and use the first part of the theorem.

Proof. If we choose $\varepsilon_1, \varepsilon_1$ to decrease slowly with n such that

1. $\lim_{n\to\infty} \varepsilon_1 = \lim_{n\to\infty} \varepsilon_1 = 0$;
2. $\lim_{n\to\infty} \varepsilon_1^2 \log n = \lim_{n\to\infty} \varepsilon_2^2 \log n = \infty$.

If so, we can derive that

$$\begin{aligned} \limsup_{n\to\infty} \frac{R_n}{\log(n)} &\leq \lim_{n\to\infty} \sum_{i:\Delta_i>0} \frac{\Delta_i}{d(\mu,\mu^*)} \frac{\log(1 + n\log^2 n)}{\log n} \\ &\leq \lim_{n\to\infty} \sum_{i:\Delta_i>0} \frac{\Delta_i}{d(\mu,\mu^*)} \frac{\log(\log^2 n + n\log^2 n)}{\log n} \\ &= \lim_{n\to\infty} \sum_{i:\Delta_i>0} \frac{\Delta_i}{d(\mu,\mu^*)} \frac{\log(n+1) + 2\log(\log n)}{\log n} \\ &= \sum_{i:\Delta_i>0} \frac{\Delta_i}{d(\mu,\mu^*)}, \end{aligned}$$

which leads to the asymptotic claim in Theorem 10.6. □

10.3 (Concentration for bounded random variables) Let $\mathbb{F} = (\mathcal{F}_t)_t$ be a filtration, $(X_t)_t$ be $[0,1]$-valued, $\mathbb{F}$-adapted sequence, such that $\mathbb{E}[X_t \mid \mathcal{F}_{t-1}] = \mu_t$ for some $\mu_1, \ldots, \mu_n \in [0,1]$ non-random numbers. Define $\mu = \frac{1}{n}\sum_{t=1}^n \mu_t$, $\hat{\mu} = \frac{1}{n}\sum_{t=1}^n X_t$. Prove that the conclusion of Lemma 10.3 still holds, i.e.,

$$\begin{aligned} &\mathbb{P}(\hat{\mu} \geq \mu + \varepsilon) \leq \exp(-nd(\mu+\varepsilon, \mu)) \text{ for } \varepsilon \in [0, 1-\mu] \ ; \\ &\mathbb{P}(\hat{\mu} \leq \mu - \varepsilon) \leq \exp(-nd(\mu-\varepsilon, \mu)) \text{ for } \varepsilon \in [0, \mu] \ . \end{aligned}$$

Hint Read Note 2 at the end of this chapter. Let $g(\cdot,\mu)$ be the cumulant-generating function of the μ-parameter Bernoulli distribution. For $X \sim \mathcal{B}(\mu)$, $\lambda \in \mathbb{R}$, $g(\lambda,\mu) = \log \mathbb{E}[\exp(\lambda X)]$. Show that $g(\lambda,\cdot)$ is concave. Next, use this and the tower rule to show that $\mathbb{E}[\exp(\lambda n(\hat{\mu}-\mu))] \le g(\lambda,\mu)^n$.

Proof. According to the Hint, let $g(\lambda,\mu) = \log \mathbb{E}\left[\exp\left(\lambda X\right)\right]$ and we would show that $g(\lambda,\cdot)$ is concave in the following. Based on the definition, we can compute that $g(\lambda,\mu) = \log(\mu\exp(\lambda) + (1-\mu)) - \lambda\mu$, which leads to $\frac{d^2}{d\mu^2} g(\lambda,\mu) = \frac{-(\exp(\lambda)-1)^2}{(\mu\exp(\lambda)+1-\mu)^2} \le 0$.

Let $S_p = \sum_{t=1}^{p}(X_t - \mu_t)$, it holds that

$$
\begin{aligned}
\mathbb{E}[\exp(\lambda S_p)] &= \mathbb{E}[\mathbb{E}[\exp(\lambda S_p)|\mathcal{F}_{p-1}]] \\
&= \mathbb{E}[\mathbb{E}[\exp(\lambda S_{p-1})\exp(\lambda(X_p - \mu_p))|\mathcal{F}_{p-1}]] \\
&= \mathbb{E}[\exp(\lambda S_{p-1})\mathbb{E}[\exp(\lambda(X_p - \mu_p))|\mathcal{F}_{p-1}]] \\
&= \mathbb{E}\left[\exp(\lambda S_{p-1})\mathbb{E}\left[\frac{\exp(\lambda X_p)}{\exp(\lambda\mu_p)}|\mathcal{F}_{p-1}\right]\right] \\
&\le \mathbb{E}\left[\exp(\lambda S_{p-1})\mathbb{E}\left[\frac{X_p(\exp(\lambda)-1)+1}{\exp(\lambda\mu_p)}|\mathcal{F}_{p-1}\right]\right] \\
&= \mathbb{E}\left[\exp(\lambda S_{p-1})\frac{\mu_p(\exp(\lambda)-1)+1}{\exp(\lambda\mu_p)}\right] \\
&= \mathbb{E}[\exp(\lambda S_{p-1})]\exp(g(\lambda,\mu_p))\,,
\end{aligned}
$$

where the inequality follows from noticing that $f(x) = \exp(\lambda x) - x(\exp(\lambda)-1) - 1 \le 0$ on $[0,1]$, as suggested in Note 2.

The above conclusion leads to $\mathbb{E}[\exp(\lambda S_n)] \le \exp\left(\sum_{t=1}^n g(\lambda,\mu_t)\right) \le \exp(ng(\lambda,\mu))$. Hence, we have

$$
\begin{aligned}
\mathbb{P}(\hat{\mu} \ge \mu + \varepsilon) &\le \frac{\mathbb{E}[\exp(\lambda\sum_{t=1}^n (X_t - \mu))]}{\exp(\lambda n\varepsilon)} \\
&\le \frac{\exp(ng(\lambda,\mu))}{\exp(\lambda n\varepsilon)} \\
&= \left(\frac{\exp(g(\lambda,\mu))}{\exp(\lambda\varepsilon)}\right)^2 \\
&= \left(\mu\exp(\lambda(1-\mu-\varepsilon)) + (1-\mu)\exp(-\lambda(\mu+\varepsilon))\right)^n\,,
\end{aligned}
$$

which can be followed by the same procedures as in the proof of Lemma 10.3, and the first inequality is due to Markov inequality that $\mathbb{P}\left(\hat{\mu} \ge \mu + \varepsilon\right) = \mathbb{P}\left(\hat{\mu} - \mu \ge \varepsilon\right) = \mathbb{P}\left(\exp(\lambda\sum_{t=1}^n (X_t - \mu)) \ge \exp(\lambda n\varepsilon)\right) \le \frac{\mathbb{E}\left[\exp(\lambda\sum_{t=1}^n (X_t - \mu))\right]}{\exp(\lambda n\varepsilon)}$. $\square$

Chapter 11 The Exp3 Algorithm

11.1 (Sampling from a multinomial) In order to implement Exp3, you need a way to sample from the exponential weights distribution. Many programming languages provide a standard way to do this. For example, in Python you can use the Numpy library and numpy.random.multinomial. In more basic languages, however, you only have access to a function rand() that returns a floating point number 'uniformly' distributed in $[0, 1]$. Describe an algorithm that takes as input a probability vector $p \in \mathcal{P}_{k-1}$ and uses a single call to rand() to return $X \in [k]$ with $P(X = i) = p_i$.

Proof. Recall that the probability vector p satisfies $\sum_{i=1}^k p_i = 1$. We can divide the interval $[0, 1]$ into several slices. For example, the first is $[0, p_1)$, the second is $[p_1, p_1 + p_2)$, ..., and the last is $[\sum_{i=1}^{k-1} p_i, 1)$. Every time we call rand(0,1) and get an output. The returned $X \in [k]$ is just the index of the slice in which the output falls. $\square$

11.2 Let π be a deterministic policy, and we define $x_{ti} = 0$ if $A_t = i$ otherwise $x_{ti} = 1$. The deterministic policy collects zero rewards all time,

$$\max_{i\in[k]} \sum_{t=1}^n x_{ti} \geq \frac{1}{k} \sum_{t=1}^n \sum_{i=1}^k x_{ti} = \frac{n(k-1)}{k}$$

11.5 Let P be a probability vector with nonzero components and let $A \sim P$. Suppose $\hat{X}$ is a function such that for all $x \in \mathbb{R}^k$,

$$\mathbb{E}\left[\hat{X}\left(A, x_A\right)\right] = \sum_{i=1}^k P_i \hat{X}\left(i, x_i\right) = x_1$$

Show that there exists an $a \in \mathbb{R}^k$ such that $< a, P >= 0$ and for all i and z in their respective domains, $\hat{X}(i, z) = a_i + \frac{\mathbb{I}\{i=1\}z}{P_1}$

Proof. Let x, x' be arbitrary but agree on the first component $x_1 = x_1'$. Let $f(x) = \sum_{i=1}^k P_i \hat{X}\left(i, x_i\right)$ Note that,

$$0 = f(x) - f(x') = \sum_{i=j}^k P_j \hat{X}\left(j, x_j\right)$$

for all $j > 1$. Since x, x' are arbitrary, $\hat{X}(j,) = \mathit{const}$. Let a_j equal to $\hat{X}(j,)$.

Further, let $a_1 = \hat{X}(1, 0)$ and then given any $x_1 \in \mathbb{R}$, $\hat{X}\left(1, x_1\right) = a_1 + x_1/P_1$.

Finally, let x be such that $x_1 = 0$. Then $0 = f(x) = \sum_i P_i a_i$. $\square$

11.6 (Variance of Exp3) In this exercise, you will show that if $\eta \in [n^{-p}, 1]$ for some $p \in (0, 1)$, then for sufficiently large n, there exists a bandit on which Exp3 has a constant probability of suffering linear regret. We work with losses so that given a bandit $y \in [0, 1]^{n\times k}$, the learner samples A_t from P_t given by

$$P_{t,i} = \frac{\exp\left(-\eta \sum_{s=1}^{t-1} \hat{Y}_{si}\right)}{\sum_{j=1}^k \exp\left(-\eta \sum_{s=1}^{t-1} \hat{Y}_{sj}\right)}$$

where $\hat{Y}_{ti} = A_{ti}y_{ti}/P_{ti}$, let $\alpha \in [1/4, 1/2]$ be a constant to be tuned subsequently and define a two-armed adversarial bandit in terms of its losses by $y_{t1} = 0$ if $t \le n/2$ and $y_{t1} = 1$ otherwise; $y_{t2} = \alpha$ if $t \le n/2$ and $y_{t2} = 0$ otherwise. For simplicity you may assume that n is even.

(a)

(b)

(c)

(d) Prove that $\mathbb{P}(R_n \ge n/4) \ge (1 - n\exp(-\eta n)/2)/65$.

(e) The previous part shows that the regret is linear with constant probability for sufficiently large n. On the other hand, a dubious application of Markov's inequality and Theorem 11.1 shows that

$$\mathbb{P}\left(\hat{R}_n \ge n/4\right) \le \frac{4\mathbb{E}\left[\hat{R}_n\right]}{n} \le O(n^{-1/2}) .$$

Explain the apparent contradiction.

Proof. (a)

(b)

(c)

(d) $\mathbb{P}\left(\hat{R}_n \ge n/4\right) \ge \mathbb{P}(A_t = 1, \forall t > n/2)$ since when $A_t = 1, \forall t > n/2$, we have $\hat{R}_n \ge n/2 - \alpha \cdot n/2 \ge n/4$. In the following, we will try to lower bound $\mathbb{P}(A_t = 1, \forall t > n/2)$.

Based on the result of (c), we fist suppose $T_2(n/2) \ge s + 1$. Then for any $t > n/2$, $\hat{L}_{t2} = \sum_{u=1}^{t-1} \hat{Y}_{u2} \ge 8\alpha n \ge 2n$. And $\hat{L}_{t1} = \sum_{u=1}^{t-1} \hat{Y}_{u1} = \sum_{u=n/2+1}^{t-1} 1/P_{u1}$. Based on the event $T_2(n/2) \ge s+1$ and the fact that $P_{t1} \ge 1/2$, we can show that $\hat{L}_{t1} \le \hat{L}_{t2}$ for all t by induction. Above all,

$$P_{t2} \le \exp\left(-\eta \sum_{s=1}^{t-1}(\hat{Y}_{s2} - \hat{Y}_{s1})\right) \le \exp\left(\eta\left(\sum_{s=1}^{t-1} \hat{Y}_{s1} - 2n\right)\right) \le \exp(-\eta n) .$$

Further, based one (c), we have $\mathbb{P}\left(\hat{R}_n \ge n/4\right) \ge \mathbb{P}(A_t = 1, \forall t > n/2) \ge 1/65\,(1 - n/2\exp(-\eta n))$.

(e) The reason is that the result obtained by using Markov's inequality is not correct since $\hat{R}_n$ may be negative.

□

11.7 First, note that if $G = -\log(-\log(U))$ then $\mathbb{P}(G \le g) = e^{-\exp(-g)}$.

$$\begin{aligned}
\mathbb{P}\left(\log a_i + G_i \ge \max_{j\in[k]} \log a_j + G_j\right) &= \mathbb{E}\left[\prod_{j\neq i} \mathbb{P}(\log a_j + G_j \le \log a_i + G_i \mid G_i)\right] \\
&= \mathbb{E}\left[\prod_{j\neq i} \exp\left(-\frac{a_j}{a_i}\exp(-G_i)\right)\right] \\
&= \mathbb{E}\left[U_i^{\sum_{j\neq i}\frac{a_j}{a_i}}\right] \\
&= \frac{1}{1 + \sum_{j\neq i}\frac{a_j}{a_i}} \\
&= \frac{a_i}{\sum_{j=1}^k a_j}
\end{aligned}$$

11.8 (Exp3 as follow-the-perturbed-leader) Let $(Z_{ti})_{ti}$ be a collection of independent and identically distributed random variables. The follow-the-perturbed-leader (FTPL) algorithm chooses

$$A_t = \operatorname*{argmax}_{i\in[k]} \left(Z_{ti} - \eta \sum_{s=1}^{t-1} \hat{Y}_{si} \right) .$$

Show that if Z_{ti} is a standard Gumbel, then follow-the-perturbed-leader is the same as Exp3.

Proof. Recall in Exp3,

$$\mathbb{P}(A_t = i) = \frac{\exp\left(\eta \sum_{s=1}^{t-1} \hat{X}_{si}\right)}{\sum_{j=1}^{k} \exp\left(\eta \sum_{s=1}^{t-1} \hat{X}_{sj}\right)} , \text{ where } \hat{X}_{ti} = 1 - \frac{\mathbb{1}\{A_t = i\}(1 - X_t)}{\mathbb{P}(A_t = i)} = 1 - Y_{ti} .$$

And in FTPL,

$$\begin{aligned}
\mathbb{P}(A_t = i) &= \mathbb{P}\left(i = \operatorname*{argmax}_{j\in[k]} Z_{tj} - \eta \sum_{s=1}^{t-1} \hat{Y}_{sj} \right) \\
&= \mathbb{P}\left(i = \operatorname*{argmax}_{j\in[k]} Z_{tj} - \eta \sum_{s=1}^{t-1} (1 - \hat{X}_{sj}) \right) \\
&= \mathbb{P}\left(i = \operatorname*{argmax}_{j\in[k]} Z_{tj} - \eta(t-1) + \eta \sum_{s=1}^{t-1} \hat{X}_{sj} \right) \\
&= \mathbb{P}\left(i = \operatorname*{argmax}_{j\in[k]} Z_{tj} + \eta \sum_{s=1}^{t-1} \hat{X}_{sj} \right) \\
&= \mathbb{P}\left(i = \operatorname*{argmax}_{j\in[k]} Z_{tj} + \log a_j \right) \\
&= \frac{a_i}{\sum_{j=1}^{k} a_j} = \frac{\exp\left(\eta \sum_{s=1}^{t-1} \hat{X}_{si}\right)}{\sum_{j=1}^{K} \exp\left(\eta \sum_{s=1}^{t-1} \hat{X}_{sj}\right)} ,
\end{aligned}$$

where we set $a_i = \exp\left(\eta \sum_{s=1}^{t-1} \hat{X}_{si}\right)$ and apply the result of 11.7.

Above all, we have shown that the policy of FTPL is equivalent to that of Exp3.

$\square$

Chapter 18

18.1

(a) By Jensen's inequality,

$$\begin{aligned}
\sum_{c\in\mathcal{C}} \sqrt{\sum_{t=1}^{n} \mathcal{I}\{c_t = c\}} &= \|C\| \sum_{c\in\mathcal{C}} \frac{1}{\|C\|} \sqrt{\sum_{t=1}^{n} \mathcal{I}\{c_t = c\}} \\
&\leq \|C\| \sqrt{\sum_{c\in\mathcal{C}} \frac{1}{\|C\|} \sum_{t=1}^{n} \mathcal{I}\{c_t = c\}} \\
&= \sqrt{\|C\| n}
\end{aligned}$$

(b) When each context occurs $\frac{n}{\|C\|}$ times we have

$$\sum_{c\in\mathcal{C}} \sqrt{\sum_{t=1}^{n} \mathcal{I}\{c_t = c\}} = \sqrt{n\|C\|}$$

Chapter 28
Follow-the-Regularised-Leader and Mirror Descent

28.5(Regret for follow-the-regularised-leader) Prove Theorem 28.5.

Theorem 2 (Theorem 28.5)**.** *((Follow-the-regularised-leader regret bound).) Let $\eta > 0$, $\mathcal{F}$ be convex with domain $\mathcal{D}$, $\mathcal{A} \subseteq \mathbb{R}^d$ be a non-empty convex set. Assume that $a_1, a_2, \ldots, a_{n+1}$ chosen by follow-the-regularised-leader are well defined. Then, for any $a \in \mathcal{A}$, the regret of follow-the-regularised-leader is bounded by*

$$R_n(a) \leq \frac{F(a) - F(a_1)}{\eta} + \sum_{t=1}^{n} \langle a_t - a_{t+1}, y_t \rangle - \frac{1}{\eta} \sum_{t=1}^{n} D\left(a_{t+1}, a_t\right) .$$

Proof. According to the definition, the regret can be decomposed as

$$R_n(a) = \sum_{t=1}^{n} \langle a_t - a, y_t \rangle = \sum_{t=1}^{n} \langle a_t - a_{t+1}, y_t \rangle + \sum_{t=1}^{n} \langle a_{t+1} - a, y_t \rangle . \tag{29}$$

Define $\Phi_t(a) = \frac{F(a)}{\eta} + \sum_{s=1}^{t} \langle a, y_s \rangle$. Then according to FTRL, $a_{t+1} = \operatorname{argmin}_{a \in \mathcal{A}} \Phi_t(a)$ for any t. The second term can then be bounded by

$$\begin{aligned}
\sum_{t=1}^{n} \langle a_{t+1} - a, y_t \rangle &= \sum_{t=1}^{n} \langle a_{t+1}, y_t \rangle - \Phi_n(a) + \frac{F(a)}{\eta} \\
&= \sum_{t=1}^{n} \left(\Phi_t(a_{t+1}) - \Phi_{t-1}(a_{t+1})\right) - \Phi_n(a) + \frac{F(a)}{\eta} \\
&= -\Phi_0(a_1) + \sum_{t=0}^{a-1} \left(\Phi_t(a_{t+1}) - \Phi_t(a_{t+2})\right) + \Phi_n(a_{n+1}) - \Phi_n(a) + \frac{F(a)}{\eta} \\
&\leq \frac{F(a) - F(a_1)}{\eta} + \sum_{t=0}^{a-1} \left(\Phi_t(a_{t+1}) - \Phi_t(a_{t+2})\right) ,
\end{aligned} \tag{30}$$

where the first two equalities are due to the definition of Φ_t and the inequality holds according to the selection rule of FTRL. For the second term, we have

$$\Phi_t(a_{t+1}) - \Phi_t(a_{t+2}) = -\langle \nabla \Phi_t(a_{t+1}), a_{t+2} - a_{t+1} \rangle - \frac{1}{\eta} D_F\left(a_{t+2}, a_{t+1}\right) \leq -\frac{1}{\eta} D_F\left(a_{t+2}, a_{t+1}\right) , \tag{31}$$

where the inequality holds according to the first-order optimality since $a_{t+1} = \operatorname{argmin}_{a \in \mathcal{A} \cap dom(F)} \Phi_t(a)$.

Above all, we can finish the proof by substituting (31) into (30) and further substituting (30) into (29). □