Worksheet 1: A Bit More Probability

Name: Due September 19, 2022

1. (Practice with Principles) Let $\mathcal{X} = \mathbb{R}$ and define $\mathbb{P}_{\mathcal{X}}(S) = \sum_{j=1}^{\infty} \frac{1}{2^j} \mathbb{1}_{r_j \in S}$ for $S \subset \mathbb{R}$, where $\{r_j\}_{j=1}^{\infty} = \mathbb{Q}$ is an enumeration of the rational numbers. Show that $\mathbb{P}_{\mathcal{X}}$ defines a probability measure. (You may want to recall rearrangement.)

2. (Function of Random Variable) Let $\mathcal{X}^2 = [0,1]^2$ with independent uniform measure $\mathbb{P}_{\mathcal{X}^2}([a,b] \times [c,d]) = (d-c)(b-a)$ for $a < b, c < d \in [0,1]$. Consider random variable $+ : \mathcal{X}^2 \to \mathbb{R}$ defined by $(x_1, x_2) \mapsto x_1 + x_2$ (see fig. 1, here $n = 2$). Express both $\mathbb{P}_{\mathcal{X}^2}(x_1 + x_2 \geq t)$ as well as the pdf (you may draw a picture).

3. (Loose Concentration) Let $(\mathcal{X} = \{0,1\}, \mathbb{P}_{\mathcal{X}}(\{1\}) = p)$ be Bernoulli p. Show that you may estimate p to precision $\varepsilon > 0$ with probability at least $1 - \delta$ by taking at least $m \geq \frac{p(1-p)}{\delta \varepsilon^2}$ samples $x_1, \ldots, x_m$. (Define random variable $s_m : \mathcal{X}^m \to \mathbb{R}$ by $(x_1, \ldots, x_m) \mapsto \frac{1}{m} \sum_{j=1}^{m} x_j$, the empirical mean, and be sure to define your measure $\mathbb{P}_{\mathcal{X}^m}$.)

4. (How fat can tails be?) For probability space $(\mathcal{X} = \mathbb{R}, \mathbb{P}_{\mathcal{X}})$ with mean 0 and unit variance, prove the following tail probability bound
$$\mathbb{P}_{\mathcal{X}}(|x| > k) \leq \frac{1}{k^2}. \tag{1}$$
Construct a measure/random variable for which $\mathbb{P}_{\mathcal{X}}(x > t) > \frac{1}{t^p} \cdot \mathbb{1}_{t>1}$ for $p \in (1,2)$, and justify why *this* bound does not contradict the one in (1).

5. (GLLN) Let $\mathcal{X} = [0,1]$ with uniform measure $\mathbb{P}_{\mathcal{X}}([a,b]) = b - a$ for $0 \leq a \leq b \leq 1$, and $\mathbb{P}_{\mathcal{X}^m} = \prod_{j=1}^{m} \mathbb{P}_{\mathcal{X}}$ be independent. Using geometric reasoning from class (again, c.f. fig. 1, here $n = m$), directly verify this instance of the Law of Large Numbers, namely that
$$\mathbb{P}_{\mathcal{X}^m}\left(\left| \frac{1}{m} \sum_{j=1}^{m} x_j - \frac{1}{2} \right| > \varepsilon \right) \xrightarrow{m \to \infty} 0,$$

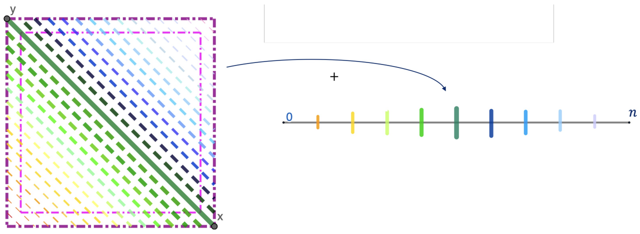

Figure 1: Geometry of LLN

give a bound on the rate of convergence, and shade in the figure region of probability concentration.

6. (Strong Concentration: Boosting) Let $\mathcal{X} = \{0,1\}$ with $\mathbb{P}_{\mathcal{X}}(\{1\}) = 1/2 + \varepsilon$ ("slightly biased coin"), and define random variable $M_m : \mathcal{X}^m \to \mathbb{R}$ by $M_m(x_1, \ldots, x_m) := \mathbb{1}_{S_m \geq 1/2}$, where $S_m := \frac{1}{m}\sum_{i=1}^{m} x_i$ denotes the empirical mean. Show that we may guarantee $M_m = 1$ with probability at least $1-\delta$ as long as $m > \frac{1}{2\varepsilon^2}\log\left(\frac{1}{\delta}\right)$. Don't forget to define your measure. (Notice that dependence on $1/\delta$ is a logarithmic relation, rather than linear as in problem #3; you will use a different concentration inequality in this problem.)

7. (Joint Measures: Conditioning and Marginalization) Consider joint probability space $\mathcal{X} \times \mathcal{Y} = \mathbb{R} \times \{0,1\}$ with measure
$$\mathbb{P}_{\mathcal{X}\times\mathcal{Y}}\big((a,b)\times\{j\}\big) = \alpha_j \cdot \int_a^b f_j(x)dx,$$
where $\alpha_0 + \alpha_1 = 1$, both nonnegative, and $f_j(x) = (p_j - 1)x^{-p_j}\mathbb{1}_{x\geq 1}$ with $p_j > 1$ (in particular, the support of each f_j is $[1,\infty)$). Express both marginals $\mathbb{P}_{\mathcal{X}}\big((a,b)\big)$ and $\mathbb{P}_{\mathcal{Y}}(y = j)$. For threshold function $\hat{y}_t : \mathcal{X} \to \mathcal{Y}$ defined by $\hat{y}_t(x) := \mathbb{1}_{x\geq t} = \begin{cases} 1 & \text{if } x \geq t \\ 0 & \text{else,} \end{cases}$ express both the true positive rate and precision of $\hat{y}_t$ defined as
$$\mathtt{tpr}(t) := \mathbb{P}_{\mathcal{X}|\mathcal{Y}}(\hat{y}_t = 1|y = 1), \text{ and } \mathtt{prec}(t) := \mathbb{P}_{\mathcal{Y}|\mathcal{X}}(y = 1|\hat{y}_t = 1).$$
For $p_0 = 5$, and $p_1 = 2$, find threshold t_o for which $\mathtt{tpr}(t_o) = 0.8$ and report what the corresponding $\mathtt{prec}(t_o)$ is. Comment on, and interpret, how precision would change if you varied (individually) α_i or p_i.