Problem 1. Let D_n be the Euclidean distance between two points, chosen uniformly at random from an n-dimensional unit hypercube, independently of each other.

(a) Determine the value of α satisfying $\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{\mathsf{D}_n}{\sqrt{n}} - \alpha\right| > \epsilon\right) = 0$ for all $\epsilon > 0$.

(b) Determine the value of β satisfying $\lim_{n\to\infty} \mathbf{P}\left(\mathsf{D}_n - \alpha\sqrt{n} \leq \gamma\right) = \Phi(\beta\gamma)$ for all $\gamma \in \mathbb{R}$ where $\Phi(\cdot)$ is standard normal CDF. (*Hint:* You can invoke the approximation $a\sqrt{n} + b \approx a\sqrt{n}$ for large n in the argument of any CDF your calculations.)

Problem 2. Let $\mathsf{X}_1, \mathsf{X}_2, \ldots$ be independent identically distributed discrete random variables with the PMF p_{X}. For any $n \in \mathbb{Z}_+$ and $\epsilon > 0$, let the typical set $\mathrm{A}_\epsilon^{(n)}$ be

$$\mathrm{A}_\epsilon^{(n)} := \left\{ x_1^n | e^{-n(H+\epsilon)} \leq p_{\mathsf{X}_1^{\mathrm{n}}}(x_1^{\mathrm{n}}) \leq e^{-n(H-\epsilon)} \right\},$$

where x_1^n and X_1^n are shorthand representations for real numbers $x_1, \ldots, x_n$ and random variables $\mathsf{X}_1, \ldots, \mathsf{X}_n$, receptively, and $H = -\sum_x p_{\mathsf{X}}(x) \ln p_{\mathsf{X}}(x)$. Show that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}\left(\mathsf{X}_1^n \notin \mathrm{A}_\epsilon^{(n)}\right) = 0$$

(*Hint*: Can you express $p_{\mathsf{X}_1^{\mathrm{n}}}(\mathsf{X}_1^{\mathrm{n}})$ in terms of the sum of independent identically distributed random variables?)

Remark. Let S_n be $\mathsf{S}_n = \sum_{t=1}^{n} \mathsf{X}_t$ for each positive integer n, where X_t are independent identically distributed random variables . One can use the central limit theorem to approximate probabilities of the form $\mathbf{P}(\mathsf{S}_n \leq \gamma)$. Questions about such probabilities are sometimes posed via random variables indexed in a different way. To familiarize yourself with such questions I recommend you read Example 5.10 and solve Problem 10 at the end of Chapter 5. The solution of Problem 10 is also presented here: https://bit.ly/2ANPxnm

Solution of Problem 1. The form of the identities used in the statement of the question suggests that the weak law of large numbers and the central limit theorem will be used in some way. In the following, we will first show that the random variables D_n can be written as a function of the sum of independent identically distributed random variables.

Let us denote the coordinates of the first point by $\mathsf{X}_1, \ldots, \mathsf{X}_n$ and the coordinates of the second point by $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$. Note that since points are chosen independently and each according to the uniform distribution on the n-dimensional unit hypercube, $\mathsf{X}_1, \ldots, \mathsf{X}_n, \mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are independent and identically distributed with a PDF that is uniform on the interval $[0, 1]$.

The Euclidean distance between the points can be written in terms of the coordinates as

$$\begin{aligned} \mathsf{D}_n &= \sqrt{\sum\nolimits_{t=1}^{n} (\mathsf{Y}_t - \mathsf{X}_t)^2} \\ &= \sqrt{\sum\nolimits_{t=1}^{n} \mathsf{Z}_t} \end{aligned} \tag{1}$$

where $\mathsf{Z}_t = (\mathsf{Y}_t - \mathsf{X}_t)^2$. Note that $\mathsf{Z}_1, \ldots, \mathsf{Z}_n$'s are independent and identically distributed because $\mathsf{X}_1, \ldots, \mathsf{X}_n, \mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are independent and identically distributed. Furthermore,

$$\begin{aligned} \mathbf{E}[\mathsf{Z_t}] &= \mathbf{E}\left[(\mathsf{Y_t} - \mathsf{X_t})^2\right] && \text{by defintion} \\ &= \mathbf{E}\left[(\mathsf{Y_t})^2\right] - 2\mathbf{E}[\mathsf{Y_t}]\,\mathbf{E}[\mathsf{X_t}] + \mathbf{E}\left[(\mathsf{X_t})^2\right] && \text{because } \mathbf{E}[\mathsf{Y_t X_t}] = \mathbf{E}[\mathsf{Y_t}]\,\mathbf{E}[\mathsf{X_t}] \text{ by the independence.} \\ &= \tfrac{1}{3} - 2\tfrac{1}{2}\tfrac{1}{2} + \tfrac{1}{3} && \text{because } \mathsf{Y}_t \text{ and } \mathsf{X}_t \text{ both unifomly distributed on } [0, 1]. \\ &= {}^1\!/_6 \end{aligned} \tag{2}$$

Thus D_n is a function of the sum of independent random variables.

(a) As a result of the weak law of large numbers and (2)

$$\lim\nolimits_{n\to\infty} \mathbf{P}\left({}^1\!/_6 - \epsilon \leq \tfrac{\sum_{t=1}^{n} \mathsf{Z}_t}{n} \leq {}^1\!/_6 + \epsilon\right) = 1 \qquad \forall \epsilon > 0$$

Using (1) we get

$$\lim\nolimits_{n\to\infty} \mathbf{P}\left(\sqrt{{}^1\!/_6 - \epsilon} \leq \tfrac{\mathsf{D}_n}{\sqrt{n}} \leq \sqrt{{}^1\!/_6 + \epsilon}\right) = 1 \qquad \forall \epsilon \in (0, \tfrac{1}{6})$$

Hence,

$$\lim\nolimits_{n\to\infty} \mathbf{P}\left(\sqrt{{}^1\!/_6} - \tilde{\epsilon} \leq \tfrac{\mathsf{D}_n}{\sqrt{n}} \leq \sqrt{{}^1\!/_6} + \tilde{\epsilon}\right) = 1 \qquad \forall \tilde{\epsilon} > 0$$

Thus $\alpha = \frac{1}{\sqrt{6}}$

(b) In order to apply the central limit theorem to $\mathsf{Z}_1, \ldots, \mathsf{Z}_n$'s are independent, we need to know the variance of Z_t's. In order to calculate that let us first calculate the second moment of Z_t's.

$$\begin{aligned} \mathbf{E}\left[\mathsf{Z_t^2}\right] &= \mathbf{E}\left[(\mathsf{Y_t} - \mathsf{X_t})^4\right] && \text{by defintion} \\ &= \sum\nolimits_{k=0}^{4} \binom{4}{k} \mathbf{E}\left[(\mathsf{Y_t})^{\mathsf{k}}\right] \mathbf{E}\left[(\mathsf{X_t})^{4-\mathsf{k}}\right] && \text{because of the independence.} \\ &= \tfrac{1}{5} - 4\tfrac{1}{4}\tfrac{1}{2} + 6\tfrac{1}{3}\tfrac{1}{3} - 4\tfrac{1}{2}\tfrac{1}{4} + \tfrac{1}{5} && \text{because } \mathsf{Y}_t \text{ and } \mathsf{X}_t \text{ both unifomly distributed on } [0, 1]. \\ &= {}^1\!/_{15} \\ var(\mathsf{Z_t}) &= \mathbf{E}\left[\mathsf{Z_t^2}\right] - \mathbf{E}[\mathsf{Z_t}]^2 \\ &= {}^1\!/_{15} - ({}^1\!/_6)^2 && \text{by (2)} \\ &= {}^7\!/_{180} \end{aligned} \tag{3}$$

Using $\alpha = {}^1\!/\sqrt{6}$ and (1) we get

$$\mathbf{P}\left(\mathsf{D}_n - \alpha\sqrt{n} \leq \gamma\right) = \mathbf{P}\left(\sum\nolimits_{t=1}^{n} \mathsf{Z}_t \leq \tfrac{n}{6} + \tfrac{\sqrt{n}}{3}\gamma + \gamma^2\right) \qquad \forall \gamma \geq -\sqrt{n/6} \tag{4}$$

On the other hand, as a result of the central limit theorem

$$\begin{aligned} \lim_{n\to\infty} \mathbf{P}\left(\sum\nolimits_{t=1}^{n} \mathsf{Z}_t \leq \tfrac{n}{6} + \tfrac{\sqrt{n}}{3}\gamma\right) &= \Phi\left(\tfrac{\gamma}{3}\tfrac{1}{\sqrt{180/7}}\right) \\ &= \Phi\left(\tfrac{\sqrt{5}}{18\sqrt{7}}\gamma\right) \end{aligned} \tag{5}$$

Note that $\frac{\sqrt{n}}{3}\gamma + \gamma^2 \approx \frac{\sqrt{n}}{3}\gamma$ for large n because $\lim\limits_{n\to\infty} \gamma^2 / \frac{\sqrt{n}}{3}\gamma = 0$. Thus using (4) and (5) we can conclude that $\beta = \frac{\sqrt{5}}{18\sqrt{7}}$.

Remark 1. Our use of the approximation $a\sqrt{n} + b \approx a\sqrt{n}$ overlooks certain subtleties, which is excusable for EE230. However, we can also rigorously prove the identity

$$\lim\nolimits_{n\to\infty} \mathbf{P}\left(\sum\nolimits_{t=1}^{n} \mathsf{Z}_t \leq \tfrac{n}{6} + \tfrac{\sqrt{n}}{3}\gamma + \gamma^2\right) = \Phi\left(\tfrac{\sqrt{5}}{18\sqrt{7}}\gamma\right). \tag{6}$$

First, note that $\mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma + \gamma^2\right) \geq \mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma\right)$ and (5) imply

$$\lim_{n\to\infty} \mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma + \gamma^2\right) \geq \Phi(\tfrac{\sqrt{5}}{18\sqrt{7}}\gamma) \tag{7}$$

if the limit exists.
Since $\mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma + \gamma^2\right) \leq \mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}(\gamma + \epsilon)\right)$ for all $n \geq \frac{9\gamma^2}{\epsilon}$, equation (5) implies

$$\lim_{n\to\infty} \mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma + \gamma^2\right) \leq \Phi(\tfrac{\sqrt{5}}{18\sqrt{7}}(\gamma + \epsilon))$$

if the limit exists. On the other $\lim_{\epsilon\to 0} \Phi(\frac{\sqrt{5}}{18\sqrt{7}}(\gamma + \epsilon)) = \Phi(\frac{\sqrt{5}}{18\sqrt{7}}\gamma)$ because the standard normal CDF Φ is a continuous function. Thus

$$\lim_{n\to\infty} \mathbf{P}\left(\sum_{t=1}^{n} \mathsf{Z}_t \leq \frac{n}{6} + \frac{\sqrt{n}}{3}\gamma + \gamma^2\right) \leq \Phi(\tfrac{\sqrt{5}}{18\sqrt{7}}\gamma) \tag{8}$$

if the limit exists. Equations (7) and (8) imply that the limit exists and (6) holds.

□

Solution of Problem 2. Note that the independence of $\mathsf{X}_1, \mathsf{X}_2, \ldots$ implies

$$\begin{aligned} p_{\mathsf{X}_1^n}(x_1^n) &= p_{\mathsf{X}_1}(x_1) \cdot p_{\mathsf{X}_2}(x_2) \cdots p_{\mathsf{X}_n}(x_n) \\ &= \prod_{t=1}^{n} p_{\mathsf{X}_t}(x_t) && \forall x_1^n. \end{aligned}$$

On the other hand $p_{\mathsf{X}_t}(x) = p_{\mathsf{X}}(x)$ for all x because $\mathsf{X}_1, \mathsf{X}_2, \ldots$ are identically distributed discrete random variables with the PMF p_{X}. Thus

$$p_{\mathsf{X}_1^n}(x_1^n) = \prod_{t=1}^{n} p_{\mathsf{X}}(x_t) \qquad \forall x_1^n.$$

If take the logarithm of both sides of the equation we can express $p_{\mathsf{X}_1^n}(x_1^n)$ in terms of a sum

$$\begin{aligned} \ln p_{\mathsf{X}_1^n}(x_1^n) &= \ln \prod_{t=1}^{n} p_{\mathsf{X}}(x_t) \\ &= \sum_{t=1}^{n} \ln p_{\mathsf{X}}(x_t) && \forall x_1^n. \end{aligned}$$

Since above relation holds for all realizations x_1^n of random variables X_1^n, we can express $\ln p_{\mathsf{X}_1^n}(\mathsf{X}_1^n)$ as follows

$$\ln p_{\mathsf{X}_1^n}(\mathsf{X}_1^n) = \sum_{t=1}^{n} \ln p_{\mathsf{X}}(\mathsf{X}_t)$$

Note that $\ln p_{\mathsf{X}}(\mathsf{X}_t)$ is a random variable for each $t \in \mathbb{Z}_+$, because $\ln p_{\mathsf{X}}(\mathsf{X}_t)$ is just a function of the random variable X_t. Thus

$$\ln p_{\mathsf{X}_1^n}(\mathsf{X}_1^n) = \sum_{t=1}^{n} \mathsf{Y}_t \tag{9}$$

where $\mathsf{Y}_t = \ln p_{\mathsf{X}}(\mathsf{X}_t)$. Note that $\mathsf{Y}_1, \mathsf{Y}_2, \ldots$ are independent and identically distributed because $\mathsf{X}_1, \mathsf{X}_2, \ldots$ are independent and identically distributed. Furthermore,

$$\begin{aligned} \mathbf{E}[\mathsf{Y}_t] &= \sum_{x} p_{\mathsf{X}_t}(x) \ln p_{\mathsf{X}}(x) \\ &= \sum_{x} p_{\mathsf{X}}(x) \ln p_{\mathsf{X}}(x) && \text{because } p_{\mathsf{X}_t} = p_{\mathsf{X}} \\ &= -H && \text{by the value } H \text{ given in the question.} \end{aligned} \tag{10}$$

The weak law of large numbers and equations (9) and (10) imply

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{\ln p_{\mathsf{X}_1^n}(\mathsf{X}_1^n) - nH}{n}\right| > \epsilon\right) = 0.$$

Then $\lim_{n\to\infty} \mathbf{P}\left(\mathsf{X}_1^n \notin \mathrm{A}_\epsilon^{(n)}\right) = 0$ holds because

$$\left|\frac{\ln p_{\mathsf{X}_1^n}(\mathsf{X}_1^n) - nH}{n}\right| \leq \epsilon \qquad \Leftrightarrow \qquad e^{-n(H+\epsilon)} \leq p_{\mathsf{X}_1^n}(\mathsf{X}_1^n) \leq e^{-n(H-\epsilon)} \qquad \Leftrightarrow \qquad \mathsf{X}_1^n \in \mathrm{A}_\epsilon^{(n)}$$

Remark 2. The quantity H is called the Entropy of the random variable X. Note that each element of the the set $\mathrm{A}_\epsilon^{(n)}$ have roughly equal probability. One can also show using the definition of the set $\mathrm{A}_\epsilon^{(n)}$ together with bounds on its probability that

$$\lim_{n\to\infty} \frac{\ln \left|\mathrm{A}_\epsilon^{(n)}\right|}{n} = H.$$

These observation constitute a special case of a more general principle: the Asymptotic equipartition property.

□