8.5. For a one-dimensional discrete random variable over a finite interval the uniform distribution maximises the entropy. In this problem it will be shown that this is a more general rule. Consider a finite region in N dimensions, $\mathcal{R}$.

(a) Assume that $\boldsymbol{X} = X_1, \dots, X_N$ is discrete valued N-dimensional random vector with probaility function $p(\boldsymbol{x})$ such that

$$\sum_{\boldsymbol{x}\in\mathcal{R}} p(\boldsymbol{x}) = 1$$
$$p(\boldsymbol{x}) = 0, \boldsymbol{x} \notin \mathcal{R}$$

where $\mathcal{R}$ has finite number of outcomes, $\sum_{\boldsymbol{x}\in\mathcal{R}} 1 = A$. Show that the uniform distribution maximises the entropy over all such distributions.

(b) Assume that $\boldsymbol{X} = X_1, \dots, X_N$ is continuous valued N-dimensional random vector with density function $f(\boldsymbol{x})$ such that

$$\int_{\mathcal{R}} f(\boldsymbol{x}) d\boldsymbol{x} = 1$$
$$f(\boldsymbol{x}) = 0, \boldsymbol{x} \notin \mathcal{R}$$

where $\mathcal{R}$ has finite volume, $\int_{\mathcal{R}} 1 d\boldsymbol{x} = V$. Show that the uniform distribution maximises the differential entropy over all such distributions.

Solution:

(a) Since $\mathcal{R}$ is finite the number of outcomes is also finite, $\sum_{\boldsymbol{x}\in\mathcal{R}} 1 = A$. Then,

$$\begin{aligned}
H(\boldsymbol{X}) - \log A &= -\sum_{\boldsymbol{x}\in\mathcal{R}} p(\boldsymbol{x}) \log p(\boldsymbol{x}) - \log A \\
&= \sum_{\boldsymbol{x}\in\mathcal{R}} p(\boldsymbol{x}) \log \frac{1/A}{p(\boldsymbol{x})} \\
&\le \sum_{\boldsymbol{x}\in\mathcal{R}} p(\boldsymbol{x}) \Big(\frac{1/A}{p(\boldsymbol{x})} - 1 \Big) \log e \\
&= \Big(\sum_{\boldsymbol{x}\in\mathcal{R}} \frac{1}{A} - \sum_{\boldsymbol{x}\in\mathcal{R}} p(\boldsymbol{x}) \Big) \log e = 0
\end{aligned}$$

where the inequality follows from the IT inequality. That means $H(\boldsymbol{X}) \le \log A$ with equality if and only if $p(\boldsymbol{x}) = 1/A, \boldsymbol{x} \in \mathcal{R}$.

(b) Since $\mathcal{R}$ is finite its volume is also finite, $\int_{\mathcal{R}} 1 d\boldsymbol{x} = V$. Then,

$$\begin{aligned}
H(\boldsymbol{X}) - \log V &= -\int_{\mathcal{R}} f(\boldsymbol{x}) \log f(\boldsymbol{x}) d\boldsymbol{x} - \log V \\
&= \int_{\mathcal{R}} f(\boldsymbol{x}) \log \frac{1/V}{f(\boldsymbol{x})} d\boldsymbol{x} \\
&\le \int_{\mathcal{R}} f(\boldsymbol{x}) \Big(\frac{1/V}{f(\boldsymbol{x})} - 1 \Big) \log e d\boldsymbol{x} \\
&= \Big(\int_{\mathcal{R}} \frac{1}{V} d\boldsymbol{x} - \int_{\mathcal{R}} f(\boldsymbol{x}) d\boldsymbol{x} \Big) \log e = 0
\end{aligned}$$

where the inequality follows from the IT inequality. That means $H(\boldsymbol{X}) \le \log V$ with equality if and only if $f(\boldsymbol{x}) = 1/V, \boldsymbol{x} \in \mathcal{R}$.