UNIVERSITÉ LAVAL

Q1 = 5/15
Q2 = 20/20
Q3 = 20/20
Q4 = 20/20
Q5 = 28/28
Q6 = 24/24
Total = 117/127

Ayoub Echchahed
(111 274 558)

Théorie de l'information
GEL-7062

Devoir 1
Résolution de Problèmes

Travail présenté à
Mr. Jean-Yves Chouinard

Faculté des sciences et de génie
Université Laval
Hiver 2022

- **<u>Problème 1 : Mesure de l'information</u>**

5/15 (revoir les expressions de H(XY), H(Y) et de H(X|Y)

Problème 1.1 : *(mesure de l'information)*

On considère ici deux expériences statistiques représentées par les variables aléatoires X et Y, pour lesquelles l'ensemble des événements de X est (x_1, x_2, x_3, x_4) et l'ensemble des événements de Y est (y_1, y_2, y_3). La matrice de probabilités conjointes $\mathbf{P} = \{p(x_i, y_j)\}_{\substack{i=1,2,3,4 \\ j=1,2,3}}$ pour ces deux expériences est :

$$\mathbf{P} = \begin{bmatrix} p(x_1,y_1) & p(x_2,y_1) & p(x_3,y_1) & p(x_4,y_1) \\ p(x_1,y_2) & p(x_2,y_2) & p(x_3,y_2) & p(x_4,y_2) \\ p(x_1,y_3) & p(x_2,y_3) & p(x_3,y_3) & p(x_4,y_3) \end{bmatrix} = \begin{bmatrix} \frac{1}{32} & \frac{3}{32} & \frac{3}{32} & \frac{3}{32} \\ \frac{3}{32} & \frac{1}{32} & \frac{1}{32} & \frac{5}{32} \\ \frac{3}{32} & \frac{1}{32} & \frac{5}{32} & \frac{3}{32} \end{bmatrix}$$

a) Quelle quantité d'information obtient-on si l'on nous informe des résultats de X et de Y ?

b) Quelle quantité d'information obtient-on si l'on nous informe des résultats de Y ?

c) Quelle quantité d'information obtient-on si l'on nous informe des résultats de X si nous savons déjà le résultat de Y ?

- Marginal distribution of X : {0.219, 0.156, 0.281, 0.344}
- Marginal distribution of Y : {0.313, 0.313, 0.375}

Définition *(information propre)* :
L'information propre associée à un message x_k est défini par :

$$\boxed{I_X(x_k) = \log_b \frac{1}{p(x_k)} = -\log_b p(x_k)}$$

$$p(X) = \sum_{j=1}^{m_y} p(X, y_j) \quad (4.13)$$

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

$$p(Y) = \sum_{i=1}^{m_x} p(x_i, Y)$$

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

a) Sum of the information content of all of the separate events

$I_{(XY)}$ = Σ − log p (x_i, y_j)
= -6 log p (3/32) + -4 log p (1/32) + -2 log p (5/32)
= 45,846 bits

b) Sum of the information content of Y only

Marginal distribution of Y : {0.313, 0.313, 0.375}

$I_{(Y)}$ = Σ − log p (y_j)
= -log p (0.313) + -log p (0.313) + -log p (0.375)
= 4.767 bits

c) Sum of the information content of X minus total mutual information

$I_{(X)}$ = Σ − log p (x_i)
= -log p (0.219) + -log p (0.156) + -log p (0.281) + -log p (0.344)
= 8.242 bits

$I_{(X\,;\,Y)}$ = Σ log ((p (x, y)) / (p (x) p (y)))
= 1/32 log ((1/32) / (0.219) (0.313)) + 3/32 log ((3/32) / (0.219) (0.313))
+ 3/32 log ((3/32) / (0.219) (0.375)) + 3/32 log ((3/32) / (0.156) (0.313))
+ 1/32 log ((1/32) / (0.156) (0.313)) + 1/32 log ((1/32) / (0.156) (0.375))
+ 3/32 log ((3/32) / (0.281) (0.313)) + 1/32 log ((1/32) / (0.281) (0.313))
+ 5/32 log ((5/32) / (0.281) (0.375)) + 3/32 log ((3/32) / (0.344) (0.313))
+ 5/32 log ((5/32) / (0.344) (0.313)) + 3/32 log ((3/32) / (0.344) (0.375))
= 0.1374 bits

$I_{(X)}$ - $I_{(X\,;\,Y)}$ = 8.242 bits - 0.1374 bits = 8.1046 bits

- **Problème 2 : Mesure de l'information de sources corrélées**

20/20

Problème 1.3 : *(mesure de l'information de sources corrélées)*

Deux sources corrélées X et Y ont la distribution conjointe suivante :

$p(x_k, y_j)$	y_1	y_2	y_3	y_4
x_1	α	β	β	β
x_2	γ	0	0	0
x_3	γ	0	0	0
x_4	γ	0	0	0

où $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$ et $\alpha + 3\beta + 3\gamma = 1$.

a) Déterminez l'entropie conjointe $H(XY)$ en fonction de α, β et γ.

b) Déterminez les entropies $H(X)$ et $H(Y)$.

c) Déterminez les entropies conditionnelles $H(X|Y)$ et $H(Y|X)$.

d) Déterminez l'information mutuelle $I(X;Y)$.

- Marginal distribution of X : {(3β + α), (γ), (γ), (γ)}
- Marginal distribution of Y : {3γ + α}, (β), (β), (β)}

L'entropie conjointe de X et de Y (incertitude sur X et sur Y est donnée par :

$$H(XY) = \mathrm{E}\left[I_{XY}(x_k, y_j)\right] = \sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\, I_{XY}(x_k, y_j)$$

$$H(XY) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log_b p(x,y) = -\sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\log_b p(x_k, y_j)$$

$$H(X) = \sum_{k=1}^{K} p(x_k)\log_b\left[\frac{1}{p(x_k)}\right] \text{ et } H(X|Y) = \sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\log_b\left[\frac{1}{p(x_k|y_j)}\right]$$

Conditional Probability Formula

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

Probability of A and B

Probability of A given B

Probability of B

a) $H(XY) = \Sigma\, p(x_k, y_j)\, I_{XY}(x_k, y_j)$

$= -\,[p(x_k, y_j) * \log p(x_k, y_j)$

5/5

$+ p(x_k, y_j) * \log p(x_k, y_j)$

$+ p(x_k, y_j) * \log p(x_k, y_j)$

$+ p(x_k, y_j) * \log p(x_k, y_j)$

$+ p(x_k, y_j) * \log p(x_k, y_j)$

$+ p(x_k, y_j) * \log p(x_k, y_j)$

$+ p(x_k, y_j) * \log p(x_k, y_j)]$

$H(XY) = -\,[\alpha * \log_2 \alpha] + 3\,[\gamma * \log_2 \gamma] + 3\,[\beta * \log_2 \beta]$

b) $H(X) = \Sigma\, p(x_k) \log[1 / p(x_k)]$

$= (3\beta + \alpha) * \log[1 / (3\beta + \alpha)] + 3\,[(\gamma) * \log[1 / (\gamma)]$

5/5

$H(Y) = \Sigma\, p(y_j) \log[1 / p(y_j)]$

$= (3\gamma + \alpha) * \log[1 / (3\gamma + \alpha)] + 3\,[(\beta) * \log[1 / (\beta)]$

c) $H(X \mid Y) = \Sigma\, p(x_k, y_j) \log[1 / p(x_k \mid y_j)]$

5/5

$= \Sigma\, p(x_k, y_j) \log[1 / (p(x_k, y_j) / p(y_j)]$

$= 3\,[\gamma \log[1 / (\gamma / (3\gamma + \alpha))] + \alpha \log[1 / (\alpha / (3\gamma + \alpha))]$

$H(Y \mid X) = \Sigma\, p(x_k, y_j) \log[1 / p(y_j \mid x_k)]$

$= \Sigma\, p(x_k, y_j) \log[1 / (p(x_k, y_j) / p(x_k)]$

$= 3\,[\beta \log[1 / (\beta / (3\beta + \alpha))] + \alpha \log[1 / (\alpha / (3\beta + \alpha))]$

d) $I(X; Y) = H(X) - H(X \mid Y)$

$= [(3\beta + \alpha) * \log[1 / (3\beta + \alpha)] + 3\,[(\gamma) * \log[1 / (\gamma)]] - [3\,[\gamma \log[1 / (\gamma / (3\gamma + \alpha))] + \alpha \log[1 / (\alpha / (3\gamma + \alpha))]]$

5/5

- **<u>Problème 3 : Mesure de l'information de sources gaussiennes conjointes</u>**

20/20

Problème 2.2 : *(mesure de l'information de sources gaussiennes conjointes)*

Soient deux variables aléatoires conjointes gaussiennes et corrélées (X, Y) de moyennes nulles $\mu_{\mathbf{X}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ et de matrice de covariance $\mathbf{K} = \begin{bmatrix} 9 & 3 \\ 3 & 9 \end{bmatrix}$.

a) Calculez les entropies différentielles $h(X)$ et $h(Y)$ en shannons ou bits ($b = 2$).

b) Calculez l'entropie différentielle conjointe $h(X, Y)$ en shannons.

c) Calculez les entropies différentielles conditionnelles $h(X|Y)$ et $h(Y|X)$ en shannons.

d) Calculez l'information mutuelle $I(X; Y)$ entre X et Y en shannons.

$\rho = 1/3$

a) $h(X) = \frac{1}{2} \log_2 (2\pi e\sigma^2)$

5/5 $= \frac{1}{2} \log_2 (2\pi e 9)$

$= 3.63$ shannons

$h(Y) = \frac{1}{2} \log_2 (2\pi e\sigma^2)$

$= \frac{1}{2} \log_2 (2\pi e 9)$

$= 3.63$ shannons

Les entropies différentielles conjointe et marginales sont alors :

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \Rightarrow h(X) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \text{ [shannons]} \\
f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{y^2}{2\sigma^2}\right)} \Rightarrow h(Y) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \text{ [shannons]} \\
f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} e^{\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma^2} - 2\rho\frac{xy}{\sigma^2} + \frac{y^2}{\sigma^2}\right)\right]} \\
&\Rightarrow h(X,Y) = \frac{1}{2}\log_2\left[(2\pi e)^2\sigma^4\left(1-\rho^2\right)\right] \text{ [shannons]}
\end{aligned}
$$

L'information mutuelle $I(X;Y)$ entre les variables aléatoires X et Y est la différence entre la somme des entropies (différentielles), $h(X)$ et $h(Y)$, et l'entropie (différentielle) conjointe $h(X,Y)$:

$$
\begin{aligned}
I(X;Y) &= h(X) + h(Y) - h(X,Y) \\
I(X;Y) &= \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) + \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) - \frac{1}{2}\log_2\left[(2\pi e)^2\sigma^4\left(1-\rho^2\right)\right] \\
I(X;Y) &= \frac{1}{2}\log_2\left(\frac{2\pi e\sigma^2 \cdot 2\pi e\sigma^2}{(2\pi e)^2\sigma^4\left(1-\rho^2\right)}\right) = \frac{1}{2}\log_2\left(\frac{1}{\left(1-\rho^2\right)}\right)
\end{aligned}
$$

$$\boxed{I(X;Y) = -\frac{1}{2}\log_2\left(1-\rho^2\right) \text{ [shannons]}}$$

b) $h(X, Y) = \frac{1}{2} \log_2 [(2\pi e)^2 \sigma^4 (1 - \rho^2)]$

$= \frac{1}{2} \log_2 [(2\pi e)^2 9^2 (1 - (1/3)^2)]$

5/5 $= 7.179$ shannons

c) $h(X \mid Y) = h(X, Y) - h(Y)$

5/5 $= 7.179 - 3.63$

$= 3.55$ shannons

$h(Y \mid X) = h(X, Y) - h(X)$

$= 7.179 - 3.63$

$= 3.55$ shannons

d) $I(X; Y) = -\frac{1}{2} \log_2 (1 - \rho^2)$

5/5 $= -\frac{1}{2} \log_2 (1 - (1/3)^2)$

$= 0.085$ shannons

- **Problème 4 : Entropies différentielle et information mutuelle distribution gaussienne conjointe**

20/20

Problème 2.3 : *(entropies différentielle et information mutuelle distribution gaussienne conjointe)*

Soit une fonction de densité de probabilité gaussienne conjointe $f_{XY}(x,y)$ de moyennes nulles $\mu_X = \mu_Y = 0$, de variances unitaires $\sigma_X^2 = \sigma_Y^2 = 1$ et de coefficient de corrélation $\rho_{XY} = 1/2$.

a) Donnez l'expression des entropies différentielles $h(X)$ et $h(Y)$ ainsi que leur valeur numérique en shannons.

b) Calculez l'entropie différentielle conjointe $h(X,Y)$ en shannons.

c) Calculez les entropies différentielles conditionnelles $h(X|Y)$ et $h(Y|X)$ en shannons.

d) Calculez l'information mutuelle $I(X;Y)$ entre X et Y en shannons.

a) $h(X) = \frac{1}{2} \log_2 (2\pi e \sigma^2)$
$= \frac{1}{2} \log_2 (2\pi e 1)$
5/5 $= 2.0471$ shannons

$h(Y) = \frac{1}{2} \log_2 (2\pi e \sigma^2)$
$= \frac{1}{2} \log_2 (2\pi e 1)$
$= 2.0471$ shannons

Les entropies différentielles conjointe et marginales sont alors :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \Rightarrow h(X) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \text{ [shannons]}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{y^2}{2\sigma^2}\right)} \Rightarrow h(Y) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \text{ [shannons]}$$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} e^{\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma^2} - 2\rho\frac{xy}{\sigma^2} + \frac{y^2}{\sigma^2}\right)\right]}$$

$$\Rightarrow h(X,Y) = \frac{1}{2}\log_2\left[(2\pi e)^2 \sigma^4 \left(1-\rho^2\right)\right] \text{ [shannons]}$$

L'information mutuelle $I(X;Y)$ entre les variables aléatoires X et Y est la différence entre la somme des entropies (différentielles), $h(X)$ et $h(Y)$, et l'entropie (différentielle) conjointe $h(X,Y)$:

$$I(X;Y) = h(X) + h(Y) - h(X,Y)$$

$$I(X;Y) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) + \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) - \frac{1}{2}\log_2\left[(2\pi e)^2\sigma^4\left(1-\rho^2\right)\right]$$

$$I(X;Y) = \frac{1}{2}\log_2\left(\frac{2\pi e\sigma^2 \cdot 2\pi e\sigma^2}{(2\pi e)^2\sigma^4\left(1-\rho^2\right)}\right) = \frac{1}{2}\log_2\left(\frac{1}{\left(1-\rho^2\right)}\right)$$

$$\boxed{I(X;Y) = -\frac{1}{2}\log_2\left(1-\rho^2\right) \text{ [shannons]}}$$

b) $h(X, Y) = \frac{1}{2} \log_2 [(2\pi e)^2 \sigma^4 (1 - \rho^2)]$
5/5 $= \frac{1}{2} \log_2 [(17.079468)^2 \, 1(1 - 0.5^2)]$
$= 3.88667$ shannons

c) $h(X \mid Y) = h(X, Y) - h(Y)$
5/5 $= 3.88667 - 2.0471$
$= 1.839572$ shannons

$h(Y \mid X) = h(X, Y) - h(X)$
$= 3.88667 - 2.0471$
$= 1.839572$ shannons

d) $I(X; Y) = - \frac{1}{2} \log_2 (1 - 0.5^2)$
$= 0.207519$ shannons

5/5

- **<u>Problème 5 : Propriété d'Équirépartition asymptotique et codage de source</u>**

28/28

Problème 3.2 : Soit une source d'information binaire sans mémoire X ayant une distribution $\mathbf{p} = [p(x_1) = 2/3 \quad p(x_2) = 1/3]$.

a) Déterminez l'entropie $H(X)$ en Sh ou bits ($b = 2$).

b) On forme des séquences $\mathbf{x}$ de longueur $n = 6$. Déterminez le nombre de séquences contenant k fois x_1 et $(n-k)$ fois x_2, pour $k = 0, 1, \ldots, 6$.

c) Déterminez la probabilité d'une séquence contenant k x_1 et $(n-k)$ x_2, pour $k = 0, 1, \ldots, 6$.

d) Déterminez l'information propre par bit d'une séquence contenant k x_1 et $(n-k)$ x_2, pour $k = 0, 1, \ldots, 6$.

e) Déterminez la probabilité de l'ensemble des séquences contenant contenant k x_1 et $(n-k)$ x_2, pour $k = 0, 1, \ldots, 6$.

f) Déterminez la probabilité $\Pr[\mathbf{x} \in \mathcal{T}_X(\delta)]$ qu'une séquence binaire $\mathbf{x}$ de longueur n soit une séquence typique avec $\delta = 0.1$.

g) Déterminez le nombre de séquences typiques $|\mathcal{T}_X(\delta)|$ avec $\delta = 0.1$.

a) H(X) = - Σ p(x_i) log $_2$ p(x_i)

H(X) = − [(2/3) log $_2$ (2/3) + (1/3) log $_2$ (1/3)]

H(X) = 0.9183 Sh (or bit) 4/4

b)
c)
d)
e)

$$\binom{N}{n} \equiv \frac{N!}{(N-n)!n!}$$

nombre n de 0s ou x_1 dans le vecteur de longueur N	nombre de séquences $\binom{N}{n}$	probabilité de chaque séquence $p(x_1)^n\, p(x_2)^{N-n}$	somme des probabilités des séquences avec n zéros $\binom{N}{n}\, p(x_1)^n\, p(x_2)^{N-n}$

4/4 4/4 4/4 4/4

k	Nombre séquences	Probabilité séquence	Information propre par bit	Probabilité Ensemble Séquence	Entropie
0	1	0.001372	1.58	0.1372	1.5849
1	6	0.002743	1.42	0.01646	1.4183
2	15	0.005487	1.25	0.0823	1.2516
3	20	0.01097	1.08	0.2194	1.0850
4	15	0.0219	0.92	0.3285	0.9188
5	6	0.0439	0.75	0.2634	0.7516
6	1	0.0878	0.58	0.0878	0.5849

f) Probability of binary sequence x of length n is a typical sequence

$\delta = 0.1$
$N = 6$
$H(X) = 0.9183$

$b^{-N[H(X)+\delta]} \leqslant p(x) \leqslant b^{-N[H(X)-\delta]}$
$2^{-6[0.9183+0.1]} \leqslant p(x) \leqslant 2^{-6[0.9183-0.1]}$
$0.01448 \leqslant p(x) \leqslant 0.033266$

Range $\{H(X) - \varepsilon; H(X) + \varepsilon\}$
Range $\{0.8183; 1.0183\}$ 4/4
Σ % (k=4) = 32,85%?

Théorème *(théorème de Shannon-McMillan pour les séquences typiques) :*
Soit une source d'information aléatoire sans mémoire X d'entropie $H(X)$, un écart δ et $\mathbf{x} = [x_{k_1,1}, \ldots, x_{k_N,N}]$ un vecteur de longueur N généré par X. On peut choisir $N \geqslant N_0$ suffisamment grand tel que l'ensemble de tous les K^N vecteurs $\{\mathbf{x}\}$ possibles puisse être partitionné en un ensemble $\mathcal{T}_X(\delta)$ des séquences typiques et un ensemble $\mathcal{T}_X^c(\delta)$ des séquences atypiques ayant les propriétés suivantes :

1. La probabilité qu'un vecteur $\mathbf{x}$ de longueur N soit atypique est bornée par :
$$\boxed{\Pr[\mathbf{x} \in \mathcal{T}_X^c(\delta)] < \varepsilon}$$
2. Si $\mathbf{x}$ est dans l'ensemble $\mathcal{T}_X(\delta)$ alors sa probabilité $p(\mathbf{x})$ est bornée par :
$$\boxed{b^{-N[H(X)+\delta]} \leqslant p(\mathbf{x}) \leqslant b^{-N[H(X)-\delta]}}$$
3. Le nombre de séquences typiques, $|\mathcal{T}_X(\delta)|$, est borné par :
$$\boxed{(1-\varepsilon)\, b^{N[H(X)-\delta]} \leqslant |\mathcal{T}_X(\delta)| \leqslant b^{N[H(X)+\delta]}}$$

La propriété d'équirépartition asymptotique *Asymptotic Equipartition Property* (AEP), l'ensemble des séquences atypiques $\in \mathcal{T}_X^c(\delta)$ est bornée par :
$$\Pr[\mathbf{x} \in \mathcal{T}_X^c(\delta)] = \Pr\left[\left|-\frac{1}{N}\log_b p(\mathbf{x}) - H(X)\right| > \delta\right]$$

g) Number of typical sequences

$\delta = 0.1$
$N = 6$
$H(X) = 0.9183$

$(1 - \varepsilon)\, b^{N[H(X)-\delta]} \leqslant |T_X(\delta)| \leqslant b^{N[H(X)+\delta]}$
$(1 - \varepsilon)\, 2^{6[0.9183-0.1]} \leqslant p(x) \leqslant 2^{6[0.9183+0.1]}$
$30{,}06 \leqslant |T_X(\delta)| \leqslant 69.06$

Range $\{H(X) - \varepsilon; H(X) + \varepsilon\}$
Range $\{0.8183; 1.0183\}$
Hence, typical set → Set of all sequences with k=**4**.
= **15 sequences**? 4/4

- **Problème 6 : Séquences typiques** 24/24

> **3.13** *Calculation of typical set.* To clarify the notion of a typical set $A_\epsilon^{(n)}$ and the smallest set of high probability $B_\delta^{(n)}$, we will calculate the set for a simple example. Consider a sequence of i.i.d. binary random variables, $X_1, X_2, \ldots, X_n$, where the probability that $X_i = 1$ is 0.6 (and therefore the probability that $X_i = 0$ is 0.4).
>
> **(a)** Calculate $H(X)$.
>
> **(b)** With $n = 25$ and $\epsilon = 0.1$, which sequences fall in the typical set $A_\epsilon^{(n)}$? What is the probability of the typical set? How many elements are there in the typical set? (This involves computation of a table of probabilities for sequences with k 1's, $0 \le k \le 25$, and finding those sequences that are in the typical set.)
>
> **(c)** How many elements are there in the smallest set that has probability 0.9?
>
> **(d)** How many elements are there in the intersection of the sets in parts (b) and (c)? What is the probability of this intersection?

a) H (X) = −0.6 log 0.6 − 0.4 log 0.4
= 0.97095 bits 5/5

b) n = 25 9/9
ε = 0.1

k	Nombre séquences	Probabilité Ensemble séquence (%)	Entropie
0	1	0.000000000	1.321928
1	25	0.000000004	1.298530
2	300	0.000000076	1.275131
3	2300	0.000000874	1.251733
4	12650	0.000007210	1.228334
5	53130	0.000045425	1.204936
6	177100	0.000227126	1.181537
7	480700	0.000924725	1.158139
8	1081575	0.003120948	1.134740
9	2042975	0.008842685	1.111342
10	3268760	0.021222445	1.087943
11	4457400	0.043409546	1.064545
12	5200300	0.075966705	1.041146
13	5200300	0.113950058	1.017748
14	4457400	0.146507217	0.994349
15	3268760	0.16115794	0.970951
16	2042975	0.15108557	0.947552
17	1081575	0.11997972	0.924154
18	480700	0.07998648	0.900755
19	177100	0.04420305	0.877357
20	53130	0.01989137	0.853958
21	12650	0.00710406	0.830560
22	2300	0.00193747	0.807161
23	300	0.00037907	0.783763
24	25	0.00004738	0.760364
25	1	0.00000284	0.736966

- B1: Sequences falling in the typical set

Range {H(X) – ε; H(X) + ε}
Computing entropy third column → Range {0.87095; 1.07095}
Hence, typical set → Set of all sequences with k **between** 11-19.

- B2: Probability of the typical set

Using the computed cumulative probabilities
Σ % (0-19) - Σ % (0-10) =
0.9706 − 0.0343 = 0.9362 or 93.62%

- B3: Elements in the typical set

$|A^{(n)}_{\varepsilon}|$ = Σ (elements k = 0 to 19) **-** Σ (elements k = 0 to 10)
33486026 − 7119516 = 26 366 510 Elements

1. We begin by showing that with high probability, the sequence is in the typical set. By the weak law of large numbers,

$$-\frac{1}{n}\log p(X^n) \to -E[\log p(X)] = H(X) \quad \text{in probability.} \tag{7.41}$$

Loi binomiale :

$$\mathcal{S}_X = \{0, 1, \ldots, n\}$$
$$p_k = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

c) 5/5

- **C1: How many elements are there in the smallest set that has probability 0.9?**

Smallest set $B^{(n)}_{\delta}$: Σ Sequences k = {25, 24, 23…} until % = 0.9
(In other words, we want the sum of the lowest entropy sequences until we reach 90%)

Σ % (25-12) = 92.22%
Σ % (25-13) = 84.62%
→ Hence sequences with k ≥ 13 & portion sequence with k = 12

Remaining probability (k = 12):
0.9 − 0.846232 = 0.053768

The number of such sequences needed to fill this probability is around:
(0.053768 / 0.07596671) * 5200300 = 3 680 691 Sequences

Smallest set with probability 0.9 = Σ Sequences (25-13) + 3 680 691
= 16 777 216 + 3 680 691
= 20 457 907 Sequences

d) 5/5

- **D1: How many elements are there in the intersection of the sets in part (b) and (c)?**

Σ Sequences (19-13) + 3 680 691 Sequences (k = 12)
Size intersection = 33 486 026 – 16 777 216 + 3 680 691
= 20 389 501 Elements

- **D2: What is the probability of this intersection?**

P (intersection) = 0.970638 − 0.153768 + 0.053768
= 0.870638 or 87.06%