CO781 / QIC 890:

Theory of Quantum Communication

Topic 3, part 1

Joint Typicality

References:

Cover & Thomas, Chapter 8

From entropy (typicality) to correlations (Joint Typicality)

Consider 2 random variables X, Y.

<u>Def</u> [Jointly typical sequence]

Given a distribution $p(xy)$, drawn iid n times

$x^n y^n$ is δ-jointly typical if

ⓐ $\left| -\frac{1}{n} \log p(x^n) - H(X) \right| \leq \delta$ (x^n typical)

ⓑ $\left| -\frac{1}{n} \log p(y^n) - H(Y) \right| \leq \delta$ (y^n typical)

ⓒ $\left| -\frac{1}{n} \log p(x^n y^n) - H(XY) \right| \leq \delta$ ($(xy)^n$ typical)

<u>Def</u> [Jointly typical set]

$A_{n,\delta} = \{ x^n y^n \in \Omega_X^n \times \Omega_Y^n : x^n y^n \text{ jointly typical} \}$

Will see: under n iid draws of p(xy), the jointly typical set occurs with prob -> 1, so we can safely use its properties ...

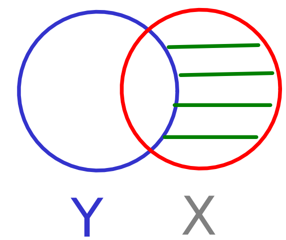
max of $p(x^n y^n)$

Obs: if $x^n y^n \in A_{n,\delta}$,

then $$p(x^n|y^n) = \frac{p(x^n y^n)}{p(y^n)} \le \frac{2^{-n(H(XY)-\delta)}}{2^{-n(H(Y)+\delta)}} = 2^{-n(H(X|Y)-2\delta)}$$

min of $p(y^n)$

Recall the chain rule
$H(XY) = \underline{H(Y)} + \underline{H(X|Y)}$.



Obs: if $x^n y^n \in A_{n,\delta}$,

then $$p(x^n|y^n) = \frac{p(x^n y^n)}{p(y^n)} \leq \frac{2^{-n(H(XY)-\delta)}}{2^{-n(H(Y)+\delta)}} \leq 2^{-n(H(X|Y)-2\delta)}$$

$$p(x^n|y^n) = \frac{p(x^n y^n)}{p(y^n)} \geq \frac{2^{-n(H(XY)+\delta)}}{2^{-n(H(Y)-\delta)}} \geq 2^{-n(H(X|Y)+2\delta)}$$

For each such y^n, the above two conditions say that

there are $\approx 2^{nH(X|Y)}$ $\tilde{x}^n$'s s.t. $\hat{x}^n y^n \in A_{n,\delta}$

Proof: similar to the proof for the AEP.

This observation is crucial for the Joint AEP ...

Thm [Joint AEP]

Using above defs, $\forall \varepsilon > 0$ $\forall \delta > 0$ $\exists n_0$ s.t. $\forall n \geq n_0$

① $\Pr(A_{n,\delta}) \geq 1-\varepsilon$

② $(1-\varepsilon)2^{n(H(XY)-\delta)} \leq |A_{n,\delta}| \leq 2^{n[H(XY)+\delta]}$

The 2 parts above are similar to the AEP.

Note that the jointly typical set is defined wrt p(xy) drawn n times iid.
Once defined, it is just a set.
Now we ask questions about the set concerning OTHER distributions.

Thm [Joint AEP]

Using above defs, $\forall \varepsilon > 0 \ \forall \delta > 0 \ \exists n_0$ s.t. $\forall n \geq n_0$

① $\Pr(A_{n,\delta}) \geq 1-\varepsilon$

② $(1-\varepsilon)\, 2^{n(H(XY)-\delta)} \leq |A_{n,\delta}| \leq 2^{n[H(XY)+\delta]}$

The 2 parts above are similar to the AEP.

$$\frac{2^{nH(XY)}}{2^{nH(X)}\ 2^{nH(Y)}}$$

③ Suppose $x^n y^n$ is drawn according to the following distribution

$$q(x^n y^n) = p(x^n) \cdot p(y^n)$$

then $2^{-n(I(X;Y)+3\delta)} \leq \Pr_q (x^n y^n \in A_{n,\delta}) \leq 2^{-n(I(X;Y)-3\delta)}$

One quick intuition to see this: there are

$2^{nH(X)}$ typical x^n's and $2^{nH(Y)}$ typical y^n's

If they're chosen indep, there are

$2^{nH(X)}\ 2^{nH(Y)}$ typical, equiprobable $x^n y^n$'s

but there are only $2^{nH(XY)}$ elements in the jointly typical set.

Thm [Joint AEP]

Using above defs, $\forall \varepsilon > 0 \;\; \forall \delta > 0 \;\; \exists n_0$ s.t. $\forall n \geq n_0$

① $\Pr(A_{n,\delta}) \geq 1-\varepsilon$

② $(1-\varepsilon)2^{n(H(XY)-\delta)} \leq |A_{n,\delta}| \leq 2^{n[H(XY)+\delta]}$

The 2 parts above are similar to the AEP.

③ Suppose $x^n y^n$ is drawn according to the following distribution

$$q(x^n y^n) = p(x^n) \cdot p(y^n)$$

then $$2^{-n(I(X:Y)+3\delta)} \leq \Pr_q(x^n y^n \in A_{n,\delta}) \leq 2^{-n(I(X:Y)-3\delta)}$$

The third part captures the "jointness" of X and Y

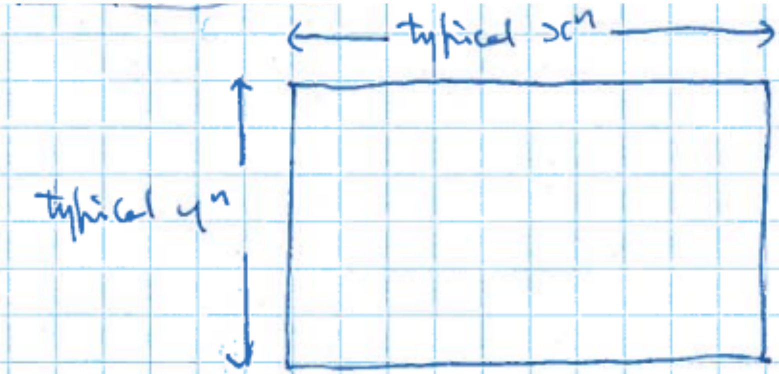
** if ** $x^n y^n$ are not from the joint distribution for XY,

but $x^n y^n$ are independent, $x^n y^n$ unlikely in the jointly typical set.

the "unlikelihood" is exp in I(X:Y)

Proof (see Cover & Thomas), based on AEP and union bound etc.

We can summarize joint typicality and JAEP by a matrix:



Total $\approx 2^{nH(X)} \times 2^{nH(Y)}$ entries

The entry associated with $x^n y^n$ is

$$\begin{cases} 1 & \text{if } x^n y^n \in A_{n,\delta} \\ 0 & \text{otherwise.} \end{cases}$$

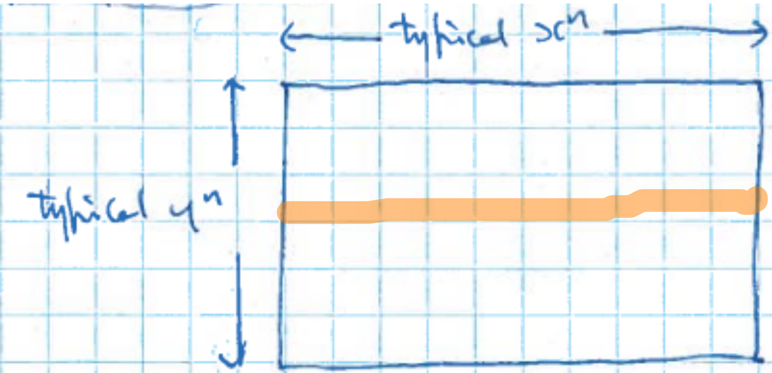
① $\Pr(A_{n,\delta}) \geq 1-\varepsilon$

② $(1-\varepsilon)2^{n(H(XY)-\delta)} \leq |A_{n,\delta}| \leq 2^{n[H(XY)+\delta]}$

① says that tolerating a prob of failure ε, we can focus on this table

② says there are $\approx 2^{nH(XY)}$ "1"'s

We can summarize joint typicality and JAEP by a matrix:



Total $\approx 2^{nH(X)} \times 2^{nH(Y)}$ entries

The entry associated with $x^n y^n$ is

$$\begin{cases} 1 & \text{if } x^n y^n \in A_{n,\delta} \\ 0 & \text{otherwise.} \end{cases}$$

① says that tolerating a prob of failure ε, we can focus on this table

② says there are $\approx 2^{nH(XY)}$ "1"'s

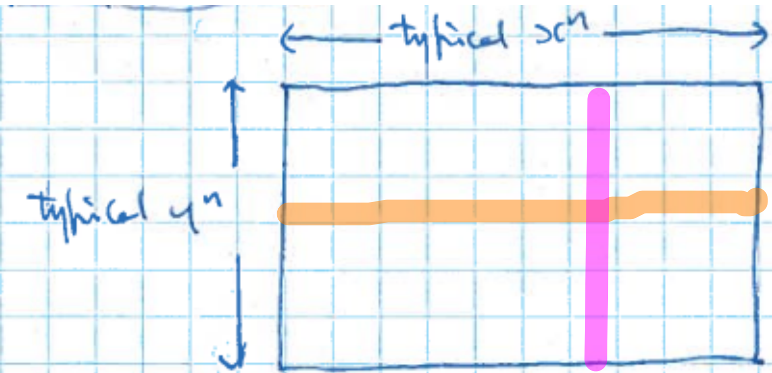
<u>Obs:</u> if $x^n y^n \in A_{n,\delta}$,

then
$$p(x^n | y^n) = \frac{p(x^n y^n)}{p(y^n)} \le \frac{2^{-n(H(XY)-\delta)}}{2^{-n(H(Y)+\delta)}} \le 2^{-n(H(X|Y)-2\delta)}$$

$$p(x^n | y^n) = \frac{p(x^n y^n)}{p(y^n)} \ge \frac{2^{-n(H(XY)+\delta)}}{2^{-n(H(Y)-\delta)}} \ge 2^{-n(H(X|Y)+2\delta)}$$

Obs says each row has approximately $2^{nH(X|Y)}$ "1"'s

We can summarize joint typicality and JAEP by a matrix:



Total $\approx 2^{nH(X)} \times 2^{nH(Y)}$ entries

The entry associated with $x^n y^n$ is

$$\begin{cases} 1 & \text{if } x^n y^n \in A_{n,\delta} \\ 0 & \text{otherwise.} \end{cases}$$

① says that tolerating a prob of failure ε, we can focus on this table

② says there are $\approx 2^{nH(XY)}$ "1"'s

<u>Obs:</u> if $x^n y^n \in A_{n,\delta}$,

then $$p(x^n|y^n) = \frac{p(x^n y^n)}{p(y^n)} \leq \frac{2^{-n(H(XY)-\delta)}}{2^{-n(H(Y)+\delta)}} \leq 2^{-n(H(X|Y)-2\delta)}$$

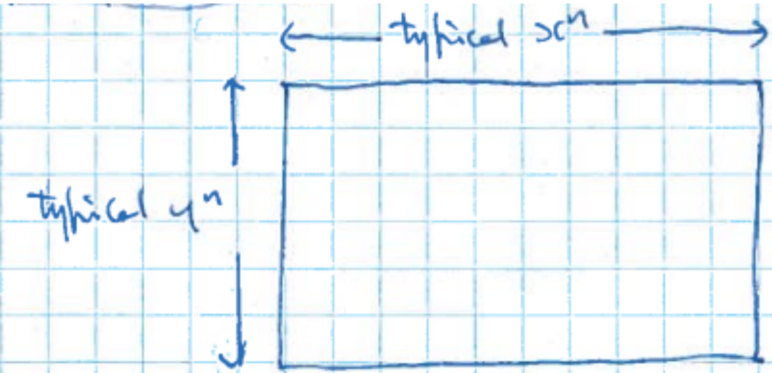
$$p(x^n|y^n) = \frac{p(x^n y^n)}{p(y^n)} \geq \frac{2^{-n(H(XY)+\delta)}}{2^{-n(H(Y)-\delta)}} \geq 2^{-n(H(X|Y)+2\delta)}$$

obs says each column has $\approx 2^{nH(Y|X)}$ "1"'s

each row . . . $2^{nH(X|Y)}$ "1"'s

XY symmetric here ...

We can summarize joint typicality and JAEP by a matrix:



Total $\approx 2^{nH(X)} \times 2^{nH(Y)}$ entries

The entry associated with $x^n y^n$ is

$$\begin{cases} 1 & \text{if } x^n y^n \in A_{n,\delta} \\ 0 & \text{otherwise.} \end{cases}$$

① says that tolerating a prob of failure ε, we can focus on this table

② says there are $\approx 2^{nH(XY)}$ "1"'s

Obs says each column has $\approx 2^{nH(Y|X)}$ "1"'s

each row . . . $2^{nH(X|Y)}$ "1"'s

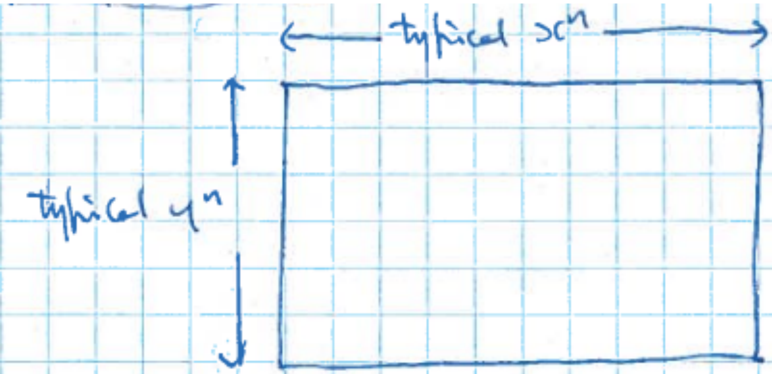
③ Suppose $x^n y^n$ is drawn according to the following distribution

$$q(x^n y^n) = p(x^n) \cdot p(y^n)$$

then $$2^{-n(I(X;Y)+3\delta)} \leq \Pr_q (x^n y^n \in A_{n,\delta}) \leq 2^{-n(I(X;Y)-3\delta)}$$

③ says a <u>random</u> entry in the table has prob $= 2^{-nI(X;Y)}$ to be "1".

We can summarize joint typicality and JAEP by a matrix:



Total $\approx 2^{nH(X)} \times 2^{nH(Y)}$ entries

The entry associated with $x^n y^n$ is

$$\begin{cases} 1 & \text{if } x^n y^n \in A_{n,\delta} \\ 0 & \text{otherwise.} \end{cases}$$

1. says that tolerating a prob of failure ε, we can focus on this table
2. says there are $\approx 2^{nH(XY)}$ "1"'s

Obs says each column has $\approx 2^{nH(Y|X)}$ "1"'s
each row $\approx 2^{nH(X|Y)}$ "1"'s

3. says a random entry in the table has prob $= 2^{-nI(X;Y)}$ to be "1".

Joint typicality gives the most critical tool to analyse classical comm through classical noisy channels. Here's an example of a joint distribution that will be relevant for this aim.

Example:

$\Omega_X = \Omega_Y = \{0,1\}$

$p(00) = \frac{1}{2}(1-e)$

$p(01) = \frac{1}{2}e$

$p(10) = \frac{1}{2}e$

$p(11) = \frac{1}{2}(1-e)$

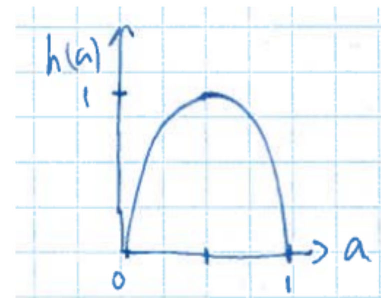
for $e \in [0,1]$,

NB. e = probability for XY to disagree, say, e = 0.1

e.g. $H(XY) = 2\left(-\frac{1}{2}\right)(1-e)\log\left(\frac{1}{2}(1-e)\right) + 2\left(-\frac{e}{2}\right)\log\left(\frac{e}{2}\right)$

$= 1 + h(e) = 1.469$

where h(a) = -a log a - (1-a) log(1-a) binary entropy function



From JAEP (1):

$\approx 2^{n\,1.469}$ jointly typical $x^n y^n$'s

For any e, H(X) = H(Y) = 1

For e = 0.1,
H(X|Y) = H(XY) - H(Y)
= 0.469

For each typical y^n, there are $\approx 2^{nH(X|Y)} = 2^{n\,0.469}$ $\tilde{x}^n$'s such that $\tilde{x}^n y^n$ is jointly typical.

For e = 0.1,
I(X:Y) = H(X) - H(X|Y)
= 0.531

For an $\tilde{x}^n$ chosen randomly for the typical set for $X_1 \ldots X_n$,

$\mathrm{Prob}(\tilde{x}^n y^n \in A_{n,\delta}) \approx \frac{2^{nH(X|Y)}}{2^{nH(X)}} \doteq 2^{-nI(X:Y)} \approx 2^{-0.531 n}$.