Subject : Intro. to Info. Theory

No. : 1

Date : 10 / 1 / 2019

* connection between

$H_\delta(X)$ and $H(X)$

- depends on δ
- can be "weak"

when $\theta = 0.25$, $H(X) = 0.81$

and $\delta = 0.25$, $H_\delta(X) = 0$

* Connection between

$\boxed{H_\delta(X^N)}$ and $H(X)^N = N \cdot H(X)$

strong(er) for large(r) N

informally : $\frac{1}{N} H_\delta(X^N) \approx H(X)$

* intuition behind the connection

sample mean $\approx$ expectation

many iid sample

$\underline{a} = (a_{i_1}, a_{i_2}, \cdots, a_{i_N})$

$$\frac{r_{i_1} + r_{i_2} + \cdots + r_{i_N}}{N} \approx \sum_i p_i r_i \quad \text{for r.v. } r_i = \log \frac{1}{p_i}$$

$$\frac{1}{N} \sum_{n=1}^{N} \log \frac{1}{p_{i_n}} \approx H(X)$$

$$= \frac{1}{N} \log \frac{1}{P(\underline{a})}$$

many $P(\underline{a}) \approx 2^{-NH}$

$\underbrace{\{ 2^{NH} \ P(\underline{a}) \}}_{\text{typical set}} \approx$ prob. 1

$NH \approx H_\delta(X^N)$

Subject : Intro. to Info. Theory

No. : 2

Date : 10 / 1 / 2019

* iid sample mean $\approx$ expectation

$$\Pr\left(\left(\underbrace{\frac{1}{N}\sum_{n=1}^{N} u_n}_{} - E(u)\right)^2 \geqslant \alpha\right) \leq \frac{\sigma^2}{\alpha N} \Rightarrow$$ weak law of large numbers

a random variable w/ mean $E(u)$

& var. $\frac{\sigma^2}{N}$ (for per-u_n variance σ^2)

$\Uparrow$

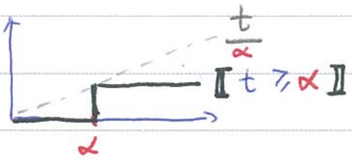
$$\Pr\left(\left(r - E(r)\right)^2 \geqslant \alpha\right) \leq \frac{\sigma_r^2}{\alpha}$$ (Chebyshef's 2nd Ineq.)

$\Uparrow$ $t: (r-E(r))^2$

$$\Pr\left(t \geqslant \alpha\right) \leq \frac{E(t)}{\alpha}$$ (Chebyshef's 1st Ineq.)

for $t \geqslant 0$

pf. $\sum p(t) [\![t \geqslant \alpha]\!] \leq \sum p(t) \frac{t}{\alpha}$



* Shannon's 1st Theorem

Given $0<\epsilon$, $0<\delta<1$, there exists N_0 s.t.

$\forall N \geqslant N_0$, $\left|\frac{1}{N} H_\delta(X^N) - H(X)\right| < \epsilon$

(a) $\frac{1}{N} H_\delta(X^N) \leq H + \epsilon$

find a "small" subset $U \subseteq X^N$

$U: \{ z : p(z) > 2^{-(H+\epsilon)\cdot N} \}$

$\Downarrow$

$|U| < 2^{(H+\epsilon)\cdot N}$

show $H_\delta(X^N) \leq \log U < (H+\epsilon)$

by showing $\Pr(U) \geqslant 1-\delta$

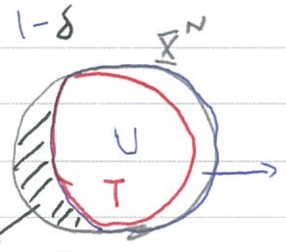
(next page)

Subject : Intro. to Info. Theory

No. : 3

Date : 10 / 1 / 2019

* $\Pr(U) \geq 1-\delta$



T_ϵ: typical set at width ϵ

$$2^{-(H+\epsilon)\cdot N} < p(z) < 2^{-(H-\epsilon)\cdot N}$$

$$\Updownarrow$$

$$-(H+\epsilon)N < \sum \log p_{i_n} < -(H-\epsilon)\cdot N$$

$$\Updownarrow$$

$$\epsilon > \left(\frac{1}{N}\sum \log \frac{1}{p_{i_n}} - H\right) > -\epsilon$$

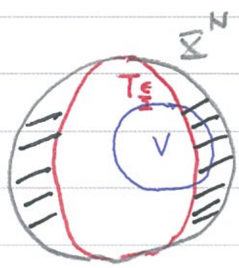
$$\Pr(U) \geq \underbrace{\Pr\left(\left(\frac{1}{N}\sum \log\frac{1}{p_{i_n}} - H\right)^2 < \epsilon^2\right)}_{T} \geq 1 - \underbrace{\frac{\sigma^2}{\epsilon^2 N}}$$

make N big so $\frac{\sigma^2}{\epsilon^2 N} \leq \delta$

* (b) $\frac{1}{N} H_\delta(\mathcal{X}^N) \geq H - \epsilon$

$$\Uparrow$$

every $V \subseteq \mathcal{X}_N$ w/ $|V| < 2^{(H-\epsilon)N}$ must have $\Pr(V) < 1-\delta$



$$\Pr(V) \leq \Pr(\text{atypical}) + |V| \cdot \underbrace{2^{-(H-\frac{\epsilon}{2})N}}_{\text{max. prob in } T_{\frac{\epsilon}{2}}}$$

$$< \frac{\sigma^2}{(\frac{\epsilon}{2})^2 N} + 2^{-\frac{\epsilon}{2}\cdot N}$$

$$< 1-\delta \quad \text{when } N \text{ big enough}$$