13.) (cont.) Kasiski Babbage

$$k \approx \frac{0{,}028433 \cdot n}{(n-1)\, I_c - 0{,}0385\, n + 0{,}06685}$$ // est. key length

use $n = 3568$, $k = 6{,}25693$

$k = 6$, done max freq. $\leq 7{,}91\%$ (first guess)

$k = 7$, done -"- $\leq 7{,}86\%$

$k = 5$, -"- $\leq 14{,}4\%$

In more detail

$$\begin{pmatrix} c_1 & c_2 \ldots & c_5 \\ c_6 & c_7 \ldots & \vdots \\ \vdots & & \vdots \\ \underbrace{c_{n-4}}_{\text{1st block}} & \ldots & \underbrace{c_n}_{\text{5th block}} \end{pmatrix}$$

Block	most freq letter char.	Freq.
1	T->E	89
2	P->E	103
3	Y->E	94
4	X->E	101
5	S->E	85

$$\begin{array}{l} \;1 \quad\;\; 2 \quad\;\; 3 \quad\;\; 4 \quad\;\; 5 \\ key = (T-E, P-E, Y-E, X-E, S-E) \end{array}$$

$= (19-4, 15-4, 24-4, 23-4, 18-4)$

$= (15, 11, 20, 19, 14)$

$= \underline{\underline{PLUTO}} = (k_0, k_1, k_2, k_3, k_4)$

Block	2nd most freq. letter char	freq	3rd most f. letter char	freq
1	I - T	68	P - A	61
2	E - T	69	T - I	56
3	N - T	63	C - I	58
4	B - T	59	G - N	53
5	H - T	68	B - N	58

frequencies in english texts

e	12,51
t	9,25
a	8,04
o	7,60
i	7,26
n	7,09

now we can decrypt with the rule

$m_i = (c_i - (K_{(i-1)} \bmod 5)) \bmod 26$ // lect p. 13

$\Rightarrow I_{\underline{m}} = 0{,}0647584 \approx K_E$

Remark to ex 11

inverse element of $A \cdot B$

$(A \cdot B)(A \cdot B)^{-1} = A \cdot \underbrace{B \cdot B^{-1}}_{E_n} \cdot A^{-1}$

since B and A are regular matrices

$= A \cdot A^{-1} = E_n$ $\square$

14.) Variance of the index of coincidence

remark 3.2. Let $Y_{ij} = \begin{cases} 1 & c_i = c_j \\ 0 & \text{otherwise} \end{cases}$

$$\begin{aligned} \text{Variance}(Y_{ij}) &= E[(Y_{ij} - E(Y_{ij}))^2] \\ &= E(Y_{ij}^2) - E^2(Y_{ij}) \\ &= E(Y_{ij}^2) - K^2 \quad // \text{ remark 3.3} \end{aligned}$$

$$\left[\begin{aligned} E(Y_{ij}^2) &= 1^2 \cdot P(Y_{ij} = 1) + 0^2 \cdot P(Y_{ij} = 0) \\ &= P(c_i = c_j) \\ &= \sum_{l=1}^{m} P(c_i = l, c_j = l) = \sum_{l=1}^{m} \underbrace{P(c_i = l)}_{q_l} \underbrace{P(c_j = l)}_{q_l} \\ &= \sum_{l=1}^{m} q_l^2 = K \end{aligned} \right]$$

Crypt1 Ü5

$$Var(Y_{i_S}) = K - K^2 = K(1-K)$$

<u>15.)</u>

a.) show for any function $f : X(\Omega) \times X(\Omega) \Rightarrow R$

$\Rightarrow H(X, Y, f(X,Y)) = H(X,Y)$

$$\bullet \Rightarrow H(X,Y,Z=f(X,Y)) \overset{\text{Remark 4.2.}}{=} \sum_{x,y,z} P(X=x, Y=y, Z=z) \cdot \log(\overbrace{X=x, Y=y, Z=z}^{P(\quad)})$$

$$P(X=y, Y=y, Z=z) = \begin{cases} P(X=x, Y=y) & \text{if } z = f(x,y) \\ 0 & \text{if } z \neq f(x,y) \end{cases}$$

$$\Rightarrow H(X,Y,Z=f(x,y)) = \sum_{x,y} P(X=x, Y=y) \log(P(X=x, Y=y))$$

$$= H(X,Y) \qquad \square \qquad // \text{with } \log(0) \cdot 0 = 0$$

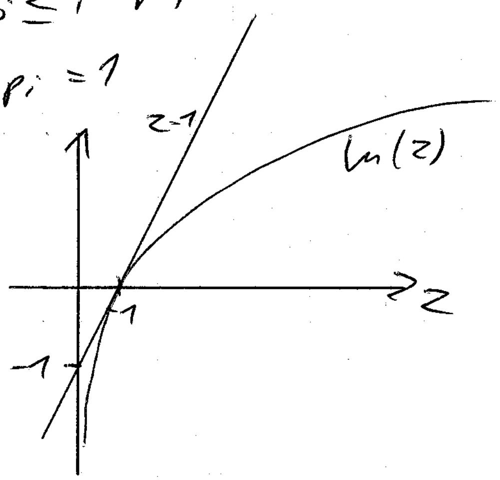
<u>16.)</u>

a.) Show that $0 \leq H(X)$

$$H(X) = -\sum_x P(X=x) \cdot \log(P(X=x))$$

$$= \sum_i^m p_i \log(p_i) \qquad 0 \leq p_i \leq 1 \ \forall i$$

$$\sum p_i = 1$$



Hint: $\ln(z) \leq z - 1$

$\Leftrightarrow -\ln(z) \geq 1 - z$

$$\geq +\sum_i \underbrace{p_i}_{\geq 0} \underbrace{(1-p_i)}_{\geq 0} \geq 0$$

b.) show that $\boxed{H(x) \le \log(m)}$

$$H(x) \le -\sum_{i=1}^{m} p_i \cdot \log(p_i) \qquad 0 \le p_i \le 1$$

$$= \sum_{i=1}^{m} p_i \cdot \log\left(\frac{1}{p_i}\right) \qquad \sum_i p_i = 1$$

//expectation value $E(x) = \sum x \cdot f(x)$

$$f(x) = \log(x)$$

$\log(x)$ is a concave function:

$$\frac{d^2}{dx^2} \log(x) = \frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2} \le 0 \qquad \forall x \ge 0$$

For concave functions: inverse Jensen inequality

$$E(f(x)) \le f(E(x))$$

$$\Rightarrow H(x) \le \log\Big(\sum \underbrace{p_i \cdot \frac{1}{p_i}}_{n}\Big) = \log(m)$$

and with $p_i = \frac{1}{m} \; \forall i \Rightarrow H(x) = \sum_{i=1}^{m} \frac{1}{m} \log(m)$

$$= \log(m) \qquad \square$$

c.) show that $H(Y|X) \le H(X)$ (conditioning reduces entropy)

$$\Rightarrow H(x) - H(X|Y) \ge 0$$

$$\Rightarrow H(x) - H(x|y) = -\sum p_i \log(p_i) + \sum_{ij} p_{ij} \log\left(\frac{p_{ij}}{p_j}\right)$$

$$= -\sum_i \sum_j p_{ij} \log(p_i) + \sum\sum p_{ij} \cdot \log\left(\frac{p_{ij}}{p_j}\right)$$

$$= -\sum\sum p_{ij} \cdot \log\left(\frac{p_i \cdot p_j}{p_{ij}} \cdot p_{ij}\right)$$

$$\ge \sum\sum p_{ij} \cdot \left(1 - \frac{p_i p_j}{p_{ij}}\right) = \sum\sum p_{ij} - \sum\sum p_{ij} \frac{p_i \cdot p_j}{p_{ij}}$$

$$= 1 - \sum\sum p_i p_j = 1 - 1 = \underline{\underline{0}} \qquad \square$$