If the cyphertext was obtained from a polyalphabetic cipher then the index of coincidence can also be used to estimate the period of the cipher.

Let p be the period of the cyphertext and place the letters of the cyphertext into groups of p so that the letters in the i^{th} position of the groups are all encrypted with the same key.

- Let $M_\alpha^{(i)}$ equal the number of occurrences of the letter α that appears in the i^{th} positions in the groups.
- If there are M groups of p, then $\sum_{\alpha=A}^{Z} M_\alpha^{(i)} = M$
- We also have $N = Mp$
- Also we can estimate that $M_\alpha^{(i)} \approx Mp_{\sigma(\alpha)}$ (again for some permutation for the alphabet σ)

english letter stats for some permutation σ of the alphabet

$$I_c = \frac{\sum_{\alpha=A} \binom{N_\alpha}{2}}{\binom{N}{2}}$$

$N_\alpha = \#$ of α which appear in the text.

Now, we calculate that $M_\alpha^{(i)} \approx M \cdot p_{\sigma(\alpha)}$

$$2D_c = \sum_{i=1}^{p} \sum_{\alpha=A}^{Z} M_\alpha^{(i)}(M_\alpha^{(i)} - 1) + 2\sum_{i=1}^{p} \sum_{j=i+1}^{p} \sum_{\alpha=A}^{Z} M_\alpha^{(i)} M_\alpha^{(j)}$$

$$\approx M^2 p(.065) - pM + M^2(.038)p(p-1)$$

$N = M \cdot p$

$$= \frac{N^2}{p}(.027) - N + N^2(.038)$$

$$\sum_{\alpha=A}^{Z} \left(M_\alpha^{(i)}\right)^2 = M^2 \sum_{\alpha=A}^{Z} p^2_{\sigma(\alpha)} \approx M^2(.065)$$

$$\sum_{\alpha=A}^{Z} M_\alpha^{(i)} = M \sum_{\alpha=A}^{Z} p_{\sigma(\alpha)} = M$$

Group 1	Group 2	Group 3	Group 4
A	R	K	L
D	U	V	M
R	L	V	H
K	W	A	Z
A	⋮		

cyphertext

put into "p" groups here $p=4$

$M_A^{(1)}$ = # of A's in column 1

$M_B^{(1)}$ = # of B's in column 1

D_c = # of equal letters Anywhere in the ~~cyph~~ cyphertext

$$= \sum_{\alpha=A}^{Z} \binom{N_\alpha}{2}$$

$$2D_c = \cancel{\sum_{\alpha=A}^{Z}} \underbrace{\sum_{i=1}^{P} \sum_{\alpha=A}^{Z} 2\cdot\binom{M_\alpha^{(i)}}{2}}_{\text{two letters in the same column}} + \underbrace{2\sum_{i=1}^{P}\sum_{j=i+1}^{P}\sum_{\alpha=A}^{Z} M_\alpha^{(i)} M_\alpha^{(j)}}_{\text{pair of letters from different columns}}$$

for each column

$$\sum_{\alpha=A}^{Z} M_\alpha^{(i)} M_\alpha^{(j)} = M^2 \sum_{\alpha=A}^{Z} P_{\sigma(\alpha)} P_{\tau(\alpha)}$$

probability that two pairs of letters are equal in random text.

$$\approx M^2 (.038)$$

Note that because $I_c = \frac{D_c}{\binom{N}{2}}$, we have that

$$2D_c = N(N-1)I_c.$$

$$I_c = \frac{D_c}{\binom{N}{2}}$$

And we just derived that

$$2D_c \approx \frac{N^2}{p}(.027) - N + N^2(.038)$$

$$2\binom{N}{2}I_c = 2D_c$$

Therefore,

$$N(N-1)I_c \approx \frac{N^2}{p}(.027) - N + N^2(.038)$$

$$(N-1)I_c \approx \frac{N}{p}(.027) - 1 + N(.038)$$

$$(N-1)I_c + 1 \approx \frac{N}{p}(.027) + N(.038)$$

$$(N-1)I_c + 1 - N(.038) \approx \frac{N}{p}(.027)$$

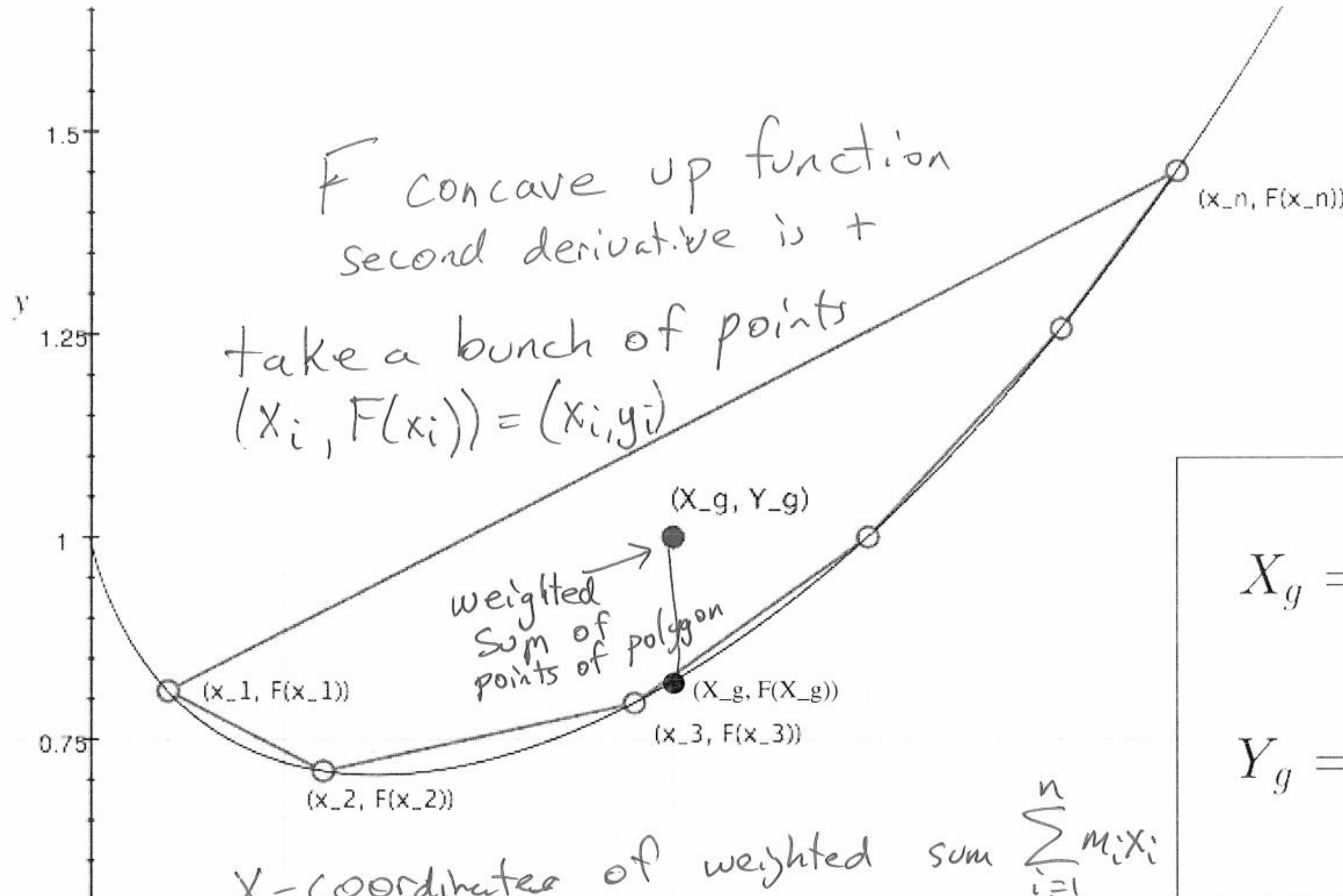
$$p((N-1)I_c + 1 - N(.038)) \approx N(.027)$$

$$p \approx \frac{N(.027)}{(N-1)I_c + 1 - N(.038)}$$

solve for p

F concave up function
second derivative is +

take a bunch of points
$(x_i, F(x_i)) = (x_i, y_i)$



weighted sum of polygon points of polygon

$$m_i \geq 0$$

$$\sum_{i=1}^{n} m_i = 1$$

$$X_g = \sum_{i=1}^{n} m_i x_i$$

$$Y_g = \sum_{i=1}^{n} m_i F(x_i)$$

X-coordinate of weighted sum $\sum_{i=1}^{n} m_i x_i$

y-coordinate of weighted sum $= \sum_{i=1}^{n} m_i y_i = \sum_{i=1}^{n} m_i F(x_i)$

F(x coordinate of weighted sum) $= \boxed{F\left(\sum_{i=1}^{n} m_i x_i\right) = F(X_g) \leq Y_g = \sum_{i=1}^{n} m_i F(x_i)}$

$$F(x) = x \log x$$

$$F'(x) = x \cdot \frac{1}{x} + \log x$$

$$F''(x) = \frac{1}{x}$$

if $x > 0$ then this function is concave up.

Take two sets of probabilities

$$\{p_i\}_{i=1}^{n} \; \& \; \{q_i\}_{i=1}^{n}$$

$$\sum p_i = 1 \qquad \sum q_i = 1$$

$$m_i = p_i \qquad x_i = q_i/p_i$$

$$F\left(\sum_{i=1}^{n} \cancel{p_i} \cdot \left(\frac{q_i}{\cancel{p_i}}\right)\right) \leq \sum_{i=1}^{n} p_i F(q_i/p_i)$$

$$F(1) \leq \sum_{i=1}^{n} \cancel{p_i} \left(\frac{q_i}{\cancel{p_i}}\right) \log(q_i/p_i)$$

$$0 \leq \sum_{i=1}^{n} q_i (\log q_i - \log p_i)$$

$$0 \leq \sum^{n} q_i \log q_i - \sum^{n} q_i \log p_i$$

Conclusion:

If I take two sets of probabilities $\{p_i\}$ & $\{q_i\}$

then $$\sum_{i=1}^{n} q_i \log p_i \leq \sum_{i=1}^{n} q_i \log q_i$$

3	1	5	2	4
R	E	C	A	L
L	H	O	W	W
E	E	N	C	R
Y	P	T	C	Y

EARLCHWL WOECERN....

E	A	R	L	C
H	W	L	W	O
E	C	E	R	N