Week 2: Cryptanalysis and Statistical Modelling

Jay Daigle

Occidental College

September 9, 2017

E	13.11%	M	2.54%	A	8.15%	N	7.10%
T	10.47%	U	2.46%	B	1.44%	O	8.00%
A	8.15%	G	1.99%	C	2.76%	P	1.98%
O	8.00%	Y	1.98%	D	3.79%	Q	0.12%
N	7.10%	P	1.98%	E	13.11%	R	6.83%
R	6.83%	W	1.54%	F	2.92%	S	6.10%
I	6.35%	B	1.44%	G	1.99%	T	10.47%
S	6.10%	V	0.92%	H	5.26%	U	2.46%
H	5.26%	K	0.42%	I	6.35%	V	0.92%
D	3.79%	X	0.17%	J	0.13%	W	1.54%
L	3.39%	J	0.13%	K	0.42%	X	0.17%
F	2.92%	Q	0.12%	L	3.39%	Y	1.98%
C	2.76%	Z	0.08%	M	2.54%	Z	0.08%

Figure: English Letter Frequencies

Jeffrey Hoffstein, Jill Catherine Pipher, and Joseph Silverman. *An introduction to mathematical cryptography*, volume 1. Springer, 2008

th	he	an	re	er	in	on	at
168	132	92	91	88	86	71	68
	nd	st	es	en	of	te	ed
	62	53	52	51	49	46	46

Most common English bigrams (frequency per 1000 words)

Jeffrey Hoffstein, Jill Catherine Pipher, and Joseph Silverman. *An introduction to mathematical cryptography*, volume 1. Springer, 2008

th	he	an	re	er	in	on	at
168	132	92	91	88	86	71	68
	nd	st	es	en	of	te	ed
	62	53	52	51	49	46	46

Most common English bigrams (frequency per 1000 words)

Jeffrey Hoffstein, Jill Catherine Pipher, and Joseph Silverman. *An introduction to mathematical cryptography*, volume 1. Springer, 2008

the	and	ing	ent	ion	her	for	tha
1.81	.73	.72	.42	.42	.36	.34	.33
	nth	int	ere	tio	ter	est	ers
	.33	.32	.31	.31	.30	.28	.28

Most common English trigrams (percentage appearance)

http://practicalcryptography.com/cryptanalysis/letter-frequencies-various-languages/english-letter-frequencies/

JNRZR BNIGI BJRGZ IZLQR OTDNJ GRIHT USDKR ZZWLG OIBTM NRGJN
IJTZJ LZISJ NRSBL QVRSI ORIQT QDEKJ JNRQW GLOFN IJTZX QLFQL
WBIMJ ITQXT HHTBL KUHQL JZKMM LZRNT OBIMI EURLW BLQZJ GKBJT
QDIQS LWJNR OLGRI EZJGK ZRBGS MJLDG IMNZT OIHRK MOSOT QHIJL
QBRJN IJJNT ZFIZL WIZTO MURZM RBTRZ ZKBNN LFRVR GIZFL KUHIM
MRIGJ LJNRB GKHRT QJRUU RBJLW JNRZI TULGI EZLUK JRUST QZLUK
EURFT JNLKJ JNRXR S

JNRZR BNIGI BJRGZ IZLQR OTDNJ GRIHT USDKR ZZWLG OIBTM NRGJN
IJTZJ LZISJ NRSBL QVRSI ORIQT QDEKJ JNRQW GLOFN IJTZX QLFQL
WBIMJ ITQXT HHTBL KUHQL JZKMM LZRNT OBIMI EURLW BLQZJ GKBJT
QDIQS LWJNR OLGRI EZJGK ZRBGS MJLDG IMNZT OIHRK MOSOT QHIJL
QBRJN IJJNT ZFIZL WIZTO MURZM RBTRZ ZKBNN LFRVR GIZFL KUHIM
MRIGJ LJNRB GKHRT QJRUU RBJLW JNRZI TULGI EZLUK JRUST QZLUK
EURFT JNLKJ JNRXR S

Letter	R	J	I	L	Z	T	N	Q	B	G	K	U	M	O	S	H	W	F	E	D	X	V
Frequency	33	30	27	35	24	20	19	16	15	15	13	12	12	10	9	8	7	6	5	5	3	2

JNRZR BNIGI BJRGZ IZLQR OTDNJ GRIHT USDKR ZZWLG OIBTM NRGJN
IJTZJ LZISJ NRSBL QVRSI ORIQT QDEKJ JNRQW GLOFN IJTZX QLFQL
WBIMJ ITQXT HHTBL KUHQL JZKMM LZRNT OBIMI EURLW BLQZJ GKBJT
QDIQS LWJNR OLGRI EZJGK ZRBGS MJLDG IMNZT OIHRK MOSOT QHIJL
QBRJN IJJNT ZFIZL WIZTO MURZM RBTRZ ZKBNN LFRVR GIZFL KUHIM
MRIGJ LJNRB GKHRT QJRUU RBJLW JNRZI TULGI EZLUK JRUST QZLUK
EURFT JNLKJ JNRXR S

Letter	R	J	I	L	Z	T	N	Q	B	G	K	U	M	O	S	H	W	F	E	D	X	V
Frequency	33	30	27	35	24	20	19	16	15	15	13	12	12	10	9	8	7	6	5	5	3	2

Bigram	JN	NR	TQ	LW	RB	RZ	JL
Frequency	11	8	6	5	5	5	5

```
theZe BhIGI BteGZ IZLQe OTDht GeIHT USDKe ZZWLG OIBTM heGth
ItTZt LZISt heSBL QVeSI OeIQT QDEKt theQW GLOFh ItTZX QLFQL
WBIMt ITQXT HHTBL KUHQL tZKMM LZehT OBIMI EUeLW BLQZt GKBtT
QDIQS LWthe OLGeI EZtGK ZeBGS MtLDG IMhZT OIHeK MOSOT QHItL
QBeth ItthT ZFIZL WIZTO MUeZM eBTeZ ZKBhh LFeVe GIZFL KUHIM
MeIGt LtheB GKHeT QteUU eBtLW theZI TULGI EZLUK teUST QZLUK
EUeFT thLKt theXe S
```

```
theZe BhaGa BteGZ aZoQe OTDht GeaHT USDKe ZZWoG OaBTM heGth
atTZt oZaSt heSBo QVeSa OeaQT QDEKt theQW GoOFh atTZX QoFQo
WBaMt aTQXT HHTBo KUHQo tZKMM oZehT OBaMa EUeoW BoQZt GKBtT
QDaQS oWthe OoGea EZtGK ZeBGS MtoDG aMhZT OaHeK MOSOT QHato
QBeth atthT ZFaZo WaZTO MUeZM eBTeZ ZKBhh oFeVe GaZFo KUHaM
MeaGt otheB GKHeT QteUU eBtoW theZa TUoGa EZoUK teUST QZoUK
EUeFT thoKt theXe S
```

```
theZe BhaGa BteGZ aZoQe OiDht GeaHi USDKe ZZWoG OaBiM heGth
atiZt oZaSt heSBo QVeSa OeaQi QDEKt theQW GoOFh atiZX QoFQo
WBaMt aiQXi HHiBo KUHQo tZKMM oZehi OBaMa EUeoW BoQZt GKBti
QDaQS oWthe OoGea EZtGK ZeBGS MtoDG aMhZi OaHeK MOSOi QHato
QBeth atthi ZFaZo WaZiO MUeZM eBieZ ZKBhh oFeVe GaZFo KUHaM
MeaGt otheB GKHei QteUU eBtoW theZa iUoGa EZoUK teUSi QZoUK
EUeFi thoKt theXe S
```

```
these BhaGa BteGs asone OiDht GeaHi USDKe ssWoG OaBiM heGth
atist osaSt heSBo nVeSa Oeani nDEKt thenW GoOFh atisX noFno
WBaMt ainXi HHiBo KUHno tsKMM osehi OBaMa EUeoW Bonst GKBti
nDanS oWthe OoGea EstGK seBGS MtoDG aMhsi OaHeK MOSOi nHato
nBeth atthi sFaso WasiO MUesM eBies sKBhh oFeVe GasFo KUHaM
MeaGt otheB GKHei nteUU eBtoW thesa iUoGa EsoUK teUSi nsoUK
EUeFi thoKt theXe S
```

these chara cters asone OiDht reaHi USDKe ssWor OaciM herth
atist osaSt heSco nVeSa Oeani nDEKt thenW roOFh atisX noFno
WcaMt ainXi HHico KUHno tsKMM osehi OcaMa EUeoW const rKcti
nDanS oWthe Oorea EstrK secrS MtoDr aMhsi OaHeK MOSOi nHato
nceth atthi sFaso WasiO MUesM ecies sKchh oFeVe rasFo KUHaM
Meart othec rKHei nteUU ectoW thesa iUora EsoUK teUSi nsoUK
EUeFi thoKt theXe S

```
these chara cters asone OiDht reaHi UyDue ssWor OaciM herth
atist osayt heyco nVeya Oeani nDEut thenW roOFh atisX noFno
WcaMt ainXi HHico uUHno tsuMM osehi OcaMa EUeoW const ructi
nDany oWthe Oorea Estru secry MtoDr aMhsi OaHeu MOyOi nHato
nceth atthi sFaso WasiO MUesM ecies suchh oFeVe rasFo uUHaM
Meart othec ruHei nteUU ectoW thesa iUora EsoUu teUyi nsoUu
EUeFi thout theXe y
```

```
these chara cters asone OiDht reaHi lyDue ssWor Oacip herth
atist osayt heyco nVeya Oeani nDEut thenW roOwh atisk nowno
Wcapt ainki HHico ulHno tsupp osehi Ocapa EleoW const ructi
nDany oWthe Oorea Estru secry ptoDr aphsi OaHeu pOyOi nHato
nceth atthi swaso WasiO plesp ecies suchh oweVe raswo ulHap
peart othec ruHei ntell ectoW thesa ilora Esolu telyi nsolu
Elewi thout theke y
```

```
these chara cters asone might reaHi lygue ssWor macip herth
atist osayt heyco nVeya meani ngbut thenW romwh atisk nowno
Wcapt ainki HHico ulHno tsupp osehi mcapa bleoW const ructi
ngany oWthe morea bstru secry ptogr aphsi maHeu pmymi nHato
nceth atthi swaso Wasim plesp ecies suchh oweVe raswo ulHap
peart othec ruHei ntell ectoW thesa ilora bsolu telyi nsolu
blewi thout theke y
```

```
these chara cters asone might readi lygue ssfor macip herth
atist osayt heyco nveya meani ngbut thenf romwh atisk nowno
fcapt ainki ddico uldno tsupp osehi mcapa bleof const ructi
ngany ofthe morea bstru secry ptogr aphsi madeu pmymi ndato
nceth atthi swaso fasim plesp ecies suchh oweve raswo uldap
peart othec rudei ntell ectof thesa ilora bsolu telyi nsolu
blewi thout theke y
```

These characters, as any one might readily guess, form a cipher—that is to say, they convey a meaning; but then from what is known of Kidd, I could not suppose him capable of constructing any of the more abstruse cryptographs. I made up my mind, at once, that this was of a simple species—such, however, as would appear to the crude intellect of the sailor, absolutely insoluble without the key.

These characters, as any one might readily guess, form a cipher—that is to say, they convey a meaning; but then from what is known of Kidd, I could not suppose him capable of constructing any of the more abstruse cryptographs. I made up my mind, at once, that this was of a simple species—such, however, as would appear to the crude intellect of the sailor, absolutely insoluble without the key.

From "The Gold-Bug" by Edgar Allen Poe

These characters, as any one might readily guess, form a cipher—that is to say, they convey a meaning; but then from what is known of Kidd, I could not suppose him capable of constructing any of the more abstruse cryptographs. I made up my mind, at once, that this was of a simple species—such, however, as would appear to the crude intellect of the sailor, absolutely insoluble without the key.

From "The Gold-Bug" by Edgar Allen Poe

Ciphertext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Plaintext	-	c	-	g	b	w	r	d	a	t	u	o	p	h	m	-	n	e	y	i	l	v	f	k	-	s

Plaintext	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
Ciphertext	I	E	B	H	R	W	D	N	T	-	X	U	O	Q	L	M	-	G	Z	J	K	V	F	-	S	-

- Message too short — at least 28 letters.

- Message too short — at least 28 letters.
- Unusual text

- Message too short — at least 28 letters.
- Unusual text
  If Youth, throughout all history, had had a champion to stand up for it; to show a doubting world that a child can think; and, possibly, do it practically, you wouldn't constantly run across folks today who claim that "a child don't know anything." A child's brain starts functioning at birth; and has, amongst its many infant convolutions, thousands of dormant atoms, into which God has put a mystic possibility for noticing an adult's act, and figuring out its purport.

- Message too short — at least 28 letters.
- Unusual text

  If Youth, throughout all history, had had a champion to stand up for it; to show a doubting world that a child can think; and, possibly, do it practically, you wouldn't constantly run across folks today who claim that "a child don't know anything." A child's brain starts functioning at birth; and has, amongst its many infant convolutions, thousands of dormant atoms, into which God has put a mystic possibility for noticing an adult's act, and figuring out its purport.

  From *Gadsby*, by Ernest Vincent Wright

- Message too short — at least 28 letters.
- Unusual text

  If Youth, throughout all history, had had a champion to stand up for it; to show a doubting world that a child can think; and, possibly, do it practically, you wouldn't constantly run across folks today who claim that "a child don't know anything." A child's brain starts functioning at birth; and has, amongst its many infant convolutions, thousands of dormant atoms, into which God has put a mystic possibility for noticing an adult's act, and figuring out its purport.

  From *Gadsby*, by Ernest Vincent Wright
- A different language

- Message too short — at least 28 letters.
- Unusual text

  If Youth, throughout all history, had had a champion to stand up for it; to show a doubting world that a child can think; and, possibly, do it practically, you wouldn't constantly run across folks today who claim that "a child don't know anything." A child's brain starts functioning at birth; and has, amongst its many infant convolutions, thousands of dormant atoms, into which God has put a mystic possibility for noticing an adult's act, and figuring out its purport.

  From *Gadsby*, by Ernest Vincent Wright
- A different language
- Not a monoalphabetic cipher

Definition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n *letters. The index of coincidence of mathbfs is denoted* $\mathsf{IndCo}(\mathbf{x})$ *and is defined to be the probability that two randomly chosen characters in the string* **s** *are identical.*

Definition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n *letters. The index of coincidence of mathbfs is denoted* $\mathrm{IndCo}(\mathbf{x})$ *and is defined to be the probability that two randomly chosen characters in the string* **s** *are identical.*

Proposition

Let $\mathbf{s} = c_1 c_2 \ldots c_n$ *be a string of* n*, and let* F_i *be the frequency with which the letter* i *appears in the string* **s**. *Then*

$$\mathrm{IndCo}(\mathbf{s}) = \frac{1}{n(n-1)} \sum_{i=0}^{25} F_i(F_i - 1). \tag{1}$$

Two important values of the index of coincidence

Two important values of the index of coincidence

Proposition

1. *If* $\mathbf{s}$ *is a string of letters generated uniformly at random, then* $\text{IndCo}(\mathbf{s}) \approx .038$.

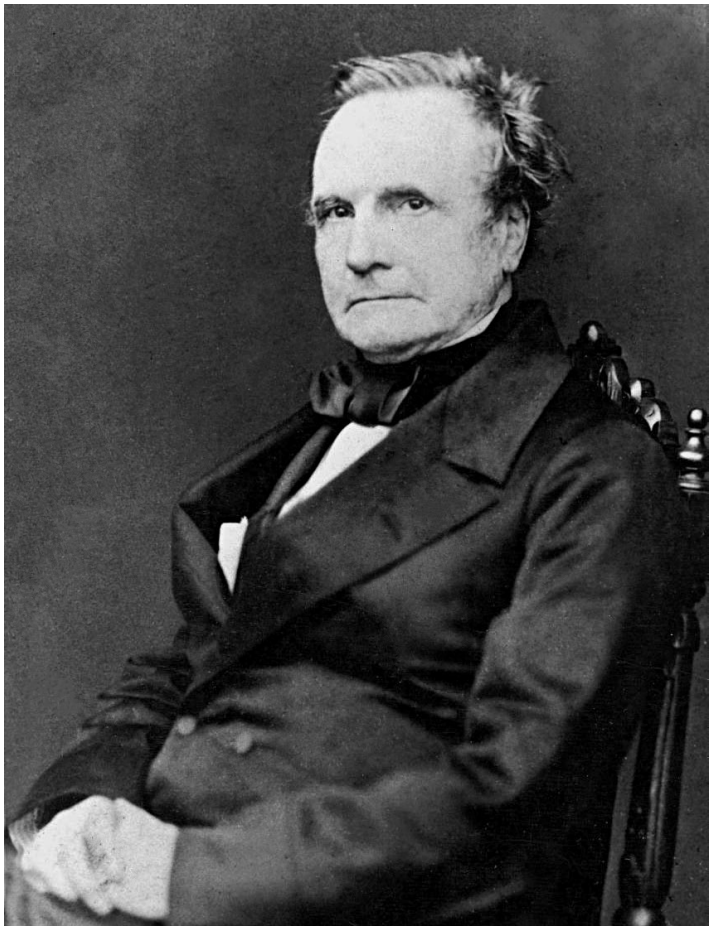
Two important values of the index of coincidence

Proposition

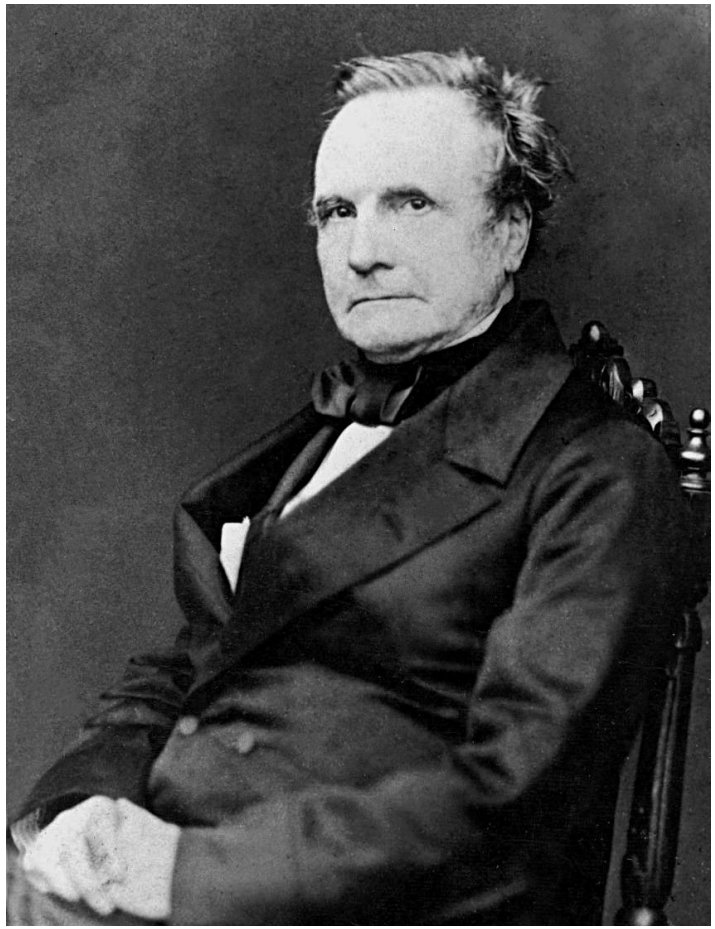
1. *If* $\mathbf{s}$ *is a string of letters generated uniformly at random, then* $\text{IndCo}(\mathbf{s}) \approx .038$.
2. *If* $\mathbf{s}$ *is a string of letters with the frequencies common in written English, then* $\text{IndCo}(\mathbf{s}) \approx .068$.

The Kasiski Method

The Kasiski Method

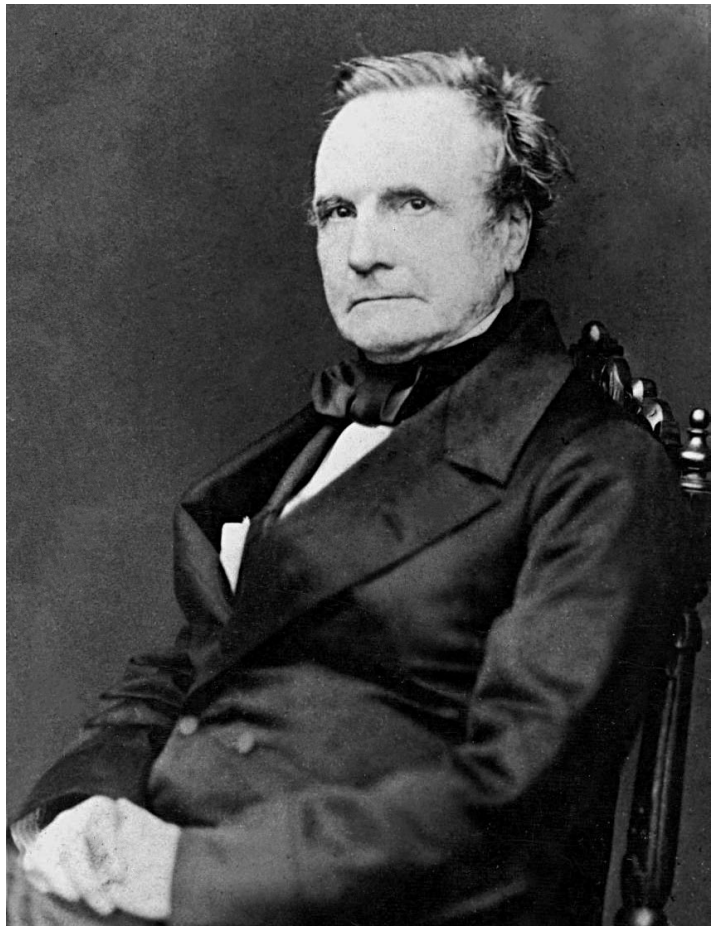


The Kasiski Method



The Kasiski method was first discovered by Charles Babbage in 1854.

The Kasiski Method



The Kasiski method was first discovered by Charles Babbage in 1854.

It was first published by Friedrich Kasiski in 1863.

```
zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
zvreg kwivs saolt nliuw oldie aqewf iiykh bjowr hdogc qhkwa
jyagg emisr zqoqh oavlk bjofr ylvps rtgiu avmsw lzgms evwpc
dmjsv jqbrn klpcf iowhv kxjbj pmfkr qthtk ozrgq ihbmq sbivd
ardym qmpbu nivxm tzwqv gefjh ucbor vwpcd xuwft qmoow jipds
fluqm oeavl jgqea lrkti wvext vkrrg xani
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
```

1 coincidence

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
3 coincidences
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
1: zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz yrigz h
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
2:  zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr zyrig zh
1 coincidence
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
3:   zp gdlrj lajkp ylxzp yyglr jgdlr zhzqy jzqre pvmsw rzyri gzh
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
4:    z pgdlr jlajk pylxz pyygl rjgdl rzhzq yjzqr epvms wrzyr
3 coincidences
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

3 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
```

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
```

5 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
```

3 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
```

2 coincidences

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
2 coincidences
```

Shift	1	2	3	4	5	6	7	8	9
Coincidences	6	6	9	5	8	13	15	11	11

```
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
5:      zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy
5 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
6:       zpgd lrjla jkpyl xzpyy glrjg dlrzh zqyjz qrepv mswrz
3 coincidences
0:zpgdl rjlaj kpylx zpyyg lrjgd lrzhz qyjzq repvm swrzy rigzh
7:        zpg dlrjl ajkpy lxzpy yglrj gdlrz hzqyj zqrep vmswr
2 coincidences
```

Shift	1	2	3	4	5	6	7	8	9
Coincidences	6	6	9	5	8	13	15	11	11

Trigrams

Trigrams

Trigram	Places	Offset	Trigram	Places	Offset
avl	117 and 258	$141 = 3 \cdot 47$	bjo	86 and 121	$35 = 5 \cdot 7$
dlr	4 and 25	$21 = 3 \cdot 7$	gdl	3 and 24	$16 = 2^4$
lrj	5 and 21	$98 = 2 \cdot 7^2$	msw	40 and 138	$84 = 2^2 \cdot 3 \cdot 7$
pcd	149 and 233	$13 = 13$	qmo	241 and 254	$98 = 2 \cdot 7^2$
vms	39 and 137	$84 = 2^2 \cdot 3 \cdot 7$	vwp	147 and 231	$84 = 2^2 3 \cdot 7$
wpc	148 and 232	$21 = 3 \cdot 7$	zhz	28 and 49	$21 = 3 \cdot 7$

Trigrams

Trigram	Places	Offset	Trigram	Places	Offset
avl	117 and 258	$141 = 3 \cdot 47$	bjo	86 and 121	$35 = 5 \cdot 7$
dlr	4 and 25	$21 = 3 \cdot 7$	gdl	3 and 24	$16 = 2^4$
lrj	5 and 21	$98 = 2 \cdot 7^2$	msw	40 and 138	$84 = 2^2 \cdot 3 \cdot 7$
pcd	149 and 233	$13 = 13$	qmo	241 and 254	$98 = 2 \cdot 7^2$
vms	39 and 137	$84 = 2^2 \cdot 3 \cdot 7$	vwp	147 and 231	$84 = 2^2 3 \cdot 7$
wpc	148 and 232	$21 = 3 \cdot 7$	zhz	28 and 49	$21 = 3 \cdot 7$

It looks like the offset is 7.

Using the Index of Coincidence

Using the Index of Coincidence

Shift	indices

Using the Index of Coincidence

Shift	indices	
2	.038	0.40

Using the Index of Coincidence

Shift	indices			
2	.038	0.40		
3	0.39	0.42	0.38	

Using the Index of Coincidence

Shift	indices				
2	.038	0.40			
3	0.39	0.42	0.38		
4	0.34	0.42	0.39	0.35	

Using the Index of Coincidence

Shift	indices				
2	.038	0.40			
3	0.39	0.42	0.38		
4	0.34	0.42	0.39	0.35	
5	0.38	0.39	0.43	0.28	0.36

Using the Index of Coincidence

Shift	indices					
2	.038	0.40				
3	0.39	0.42	0.38			
4	0.34	0.42	0.39	0.35		
5	0.38	0.39	0.43	0.28	0.36	
6	0.38	0.40	0.39	0.38	0.32	0.33

Using the Index of Coincidence

Shift	indices						
2	.038	0.40					
3	0.39	0.42	0.38				
4	0.34	0.42	0.39	0.35			
5	0.38	0.39	0.43	0.28	0.36		
6	0.38	0.40	0.39	0.38	0.32	0.33	
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64

Using the Index of Coincidence

Shift	indices							
2	.038	0.40						
3	0.39	0.42	0.38					
4	0.34	0.42	0.39	0.35				
5	0.38	0.39	0.43	0.28	0.36			
6	0.38	0.40	0.39	0.38	0.32	0.33		
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64	
8	0.37	0.29	0.38	0.33	0.34	0.57	0.40	0.39

Using the Index of Coincidence

Shift	indices							
2	.038	0.40						
3	0.39	0.42	0.38					
4	0.34	0.42	0.39	0.35				
5	0.38	0.39	0.43	0.28	0.36			
6	0.38	0.40	0.39	0.38	0.32	0.33		
7	0.62	0.57	0.65	0.60	0.60	0.64	0.64	
8	0.37	0.29	0.38	0.33	0.34	0.57	0.40	0.39

Frequency Counts on Substrings

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Frequency Counts on Substrings

`zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x`

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Or

Ciphertext	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Frequency Counts on Substrings

```
zlxrh rrhwl oehdw eokli lwvlh phqby nwhwf julrx x
```

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequency	0	1	0	1	2	1	0	6	1	1	1	6	0	1	2	1	1	4	0	0	1	1	5	3	1	1

Ciphertext	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Or

Ciphertext	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Mutual Index of Coincidence

Definition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters. Then we define the mutual index of coincidence to be* $\mathrm{MutIndCo}(\mathbf{s}, \mathbf{t})$*, the chance that a randomly selected letter of* $\mathbf{s}$ *is the same as a randomly selected letter of* $\mathbf{t}$*.*

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$
2. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from the same distribution, given by taking English frequencies and permuting the letters, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .068$.

Mutual Index of Coincidence

Proposition

Let $\mathbf{s} = c_1 c_2 \dots c_n, \mathbf{t} = d_1 d_2 \dots d_m$ *be two strings of letters, and let* $F_i(\mathbf{s})$ *be the number of times the* i*th letter appears in the string* $\mathbf{s}$*. Then:*

1. $$\text{MutIndCo}(\mathbf{s}, \mathbf{t}) = \frac{1}{nm} \sum_{i=0}^{25} F_i(\mathbf{s}) F_i(\mathbf{t}).$$
2. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from the same distribution, given by taking English frequencies and permuting the letters, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .068$.
3. *If the letters of* $\mathbf{s}$ *and* $\mathbf{t}$ *are drawn from different such distributions, then* $\text{MutIndCo}(\mathbf{s}, \mathbf{t}) \approx .038$.

i	j	σ	MutIndCo$(i, j+\sigma)$	Relative shift equation
1	3	1	.067	$\beta_1 - \beta_3 = 1$
3	7	10	.069	$\beta_3 - \beta_7 = 10$
1	4	19	.071	$\beta_1 - \beta_4 = 19$
1	6	16	.071	$\beta_1 - \beta_6 = 16$
3	4	18	.073	$\beta_3 - \beta_4 = 18$
3	5	24	.067	$\beta_3 - \beta_5 = 24$
3	6	15	.074	$\beta_3 - \beta_6 = 15$
4	6	23	.066	$\beta_4 - \beta_6 = 23$
4	7	18	.071	$\beta_4 - \beta_7 = 18$
6	7	21	.069	$\beta_6 - \beta_7 = 21$

$$\beta_3 = \beta_1 + 25$$
$$\beta_6 = \beta_1 + 10$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$
$$\beta_4 = \beta_1 + 7$$
$$\beta_7 = \beta_3 + 16 = \beta_1 + 15$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\text{Key} = \text{AFZHBKP}$$

$$\beta_3 = \beta_1 + 25 \qquad \beta_4 = \beta_1 + 7$$
$$\beta_6 = \beta_1 + 10 \qquad \beta_7 = \beta_3 + 16 = \beta_1 + 15$$
$$\beta_5 = \beta_3 + 2 = \beta_1 + 1$$

$$\text{MutIndCo}(2, 4 + 24) = .061 \Rightarrow \beta_2 = \beta_4 + 24 = \beta_1 + 5$$

$$\text{Key} = \text{AFZHBKP} + \text{shift}$$

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`
2	CHBJDMR	`xifuifsjtibmmuvso`

β_1	Keyword	Potential plaintext
0	AFZHBKP	`zkhwkhulvkdoowxuq`
1	BGAICLQ	`yjgvjgtkujcnnvwtp`
2	CHBJDMR	`xifuifsjtibmmuvso`
3	DICKENS	`whetherishallturn`

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl
6	GLFNHQV	tebqebofpexiiqrok

β_1	Keyword	Potential plaintext
0	AFZHBKP	zkhwkhulvkdoowxuq
1	BGAICLQ	yjgvjgtkujcnnvwtp
2	CHBJDMR	xifuifsjtibmmuvso
3	DICKENS	whetherishallturn
4	EJDLFOT	vgdsgdqhrgzkkstqm
5	FKEMGPU	ufcrfcpgqfyjjrspl
6	GLFNHQV	tebqebofpexiiqrok

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

"Whether I shall turn out to be the hero of my own life, or whether that station will be held by anybody else, these pages must show. To begin my life with the beginning of my life, I record that I was born (as I have been informed and believe) on a Friday, at twelve oclock at night. It was remarked that the clock began to strike, and I began to cry, simultaneously."

```
wheth erish alltu rnout tobet heher oofmy ownli feorw hethe
rthat stati onwil lbehe ldbya nybod yelse these pages musts
howto begin mylif ewith thebe ginni ngofm ylife ireco rdtha
tiwas borna sihav ebeen infor medan dbeli eveon afrid ayatt
welve ocloc katni ghtit wasre marke dthat thecl ockbe ganto
strik eandi began tocry simul taneo usly
```

"Whether I shall turn out to be the hero of my own life, or whether that station will be held by anybody else, these pages must show. To begin my life with the beginning of my life, I record that I was born (as I have been informed and believe) on a Friday, at twelve oclock at night. It was remarked that the clock began to strike, and I began to cry, simultaneously."

From *David Copperfield*, by Charles Dickens