

Figure 4. The record of 10,000 tosses of an ideal coin (described in section 6).

TABLE 3.26 Probabilities of occurrence of the letters in the English alphabet in the U.S. Constitution.

Letter	Probability	Letter	Probability
A	0.057305	N	0.056035
B	0.014876	O	0.058215
C	0.025775	P	0.021034
D	0.026811	Q	0.000973
E	0.112578	R	0.048819
F	0.022875	S	0.060289
G	0.009523	T	0.078085
H	0.042915	U	0.018474
I	0.053475	V	0.009882
J	0.002031	W	0.007576
K	0.001016	X	0.002264
L	0.031403	Y	0.011702
M	0.015892	Z	0.001502

TABLE 3.27 Probabilities of occurrence of the letters in the English alphabet in this chapter.

Letter	Probability	Letter	Probability
A	0.049855	N	0.048039
B	0.016100	O	0.050642
C	0.025835	P	0.015007
D	0.030232	Q	0.001509
E	0.097434	R	0.040492
F	0.019754	S	0.042657
G	0.012053	T	0.061142
H	0.035723	U	0.015794
I	0.048783	V	0.004988
J	0.000394	W	0.012207
K	0.002450	X	0.003413
L	0.025835	Y	0.008466
M	0.016494	Z	0.001050

Here are some examples of Markov approximations to English from Shannon's original paper [138]:

1. *Zero-order approximation.* (The symbols are independent and equiprobable.)

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ

FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD

2. *First-order approximation.* (The symbols are independent. Frequency of letters matches English text.)

OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI

ALHENHTTPA OOBTTVA NAH BRL

3. *Second-order approximation.* (The frequency of pairs of letters matches English text.)

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY

ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO

TIZIN ANDY TOBE SEACE CTISBE

4. *Third-order approximation.* (The frequency of triplets of letters matches English text.)

IN NO IST LAT WHEY CRATICT FROURE BERS GROCID

PONDENOME OF DEMONSTURES OF THE REPTAGIN IS

REGOACTIONA OF CRE

5. *Fourth-order approximation.* (The frequency of quadruplets of letters matches English text. Each letter depends on the previous three letters. This sentence is from Lucky's book, *Silicon Dreams* [183].)

THE GENERATED JOB PROVIDUAL BETTER TRAND THE
DISPLAYED CODE, ABOVERY UPONDULTS WELL THE
CODERST IN THESTICAL IT DO HOCK BOTHE MERG.
(INSTATES CONS ERATION. NEVER ANY OF PUBLE AND TO
THEORY. EVENTIAL CALLEGAND TO ELAST BENERATED IN
WITH PIES AS IS WITH THE)

Instead of continuing with the letter models, we jump to word models.

6. *First-order word model.* (The words are chosen independently but with frequencies as in English.)

REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME
CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO
OF TO EXPERT GRAY COME TO FURNISHES THE LINE
MESSAGE HAD BE THESE.

7. *Second-order word model.* (The word transition probabilities match English text.)

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH
WRITER THAT THE CHARACTER OF THIS POINT IS
THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE
TIME OF WHO EVER TOLD THE PROBLEM FOR AN
UNEXPECTED

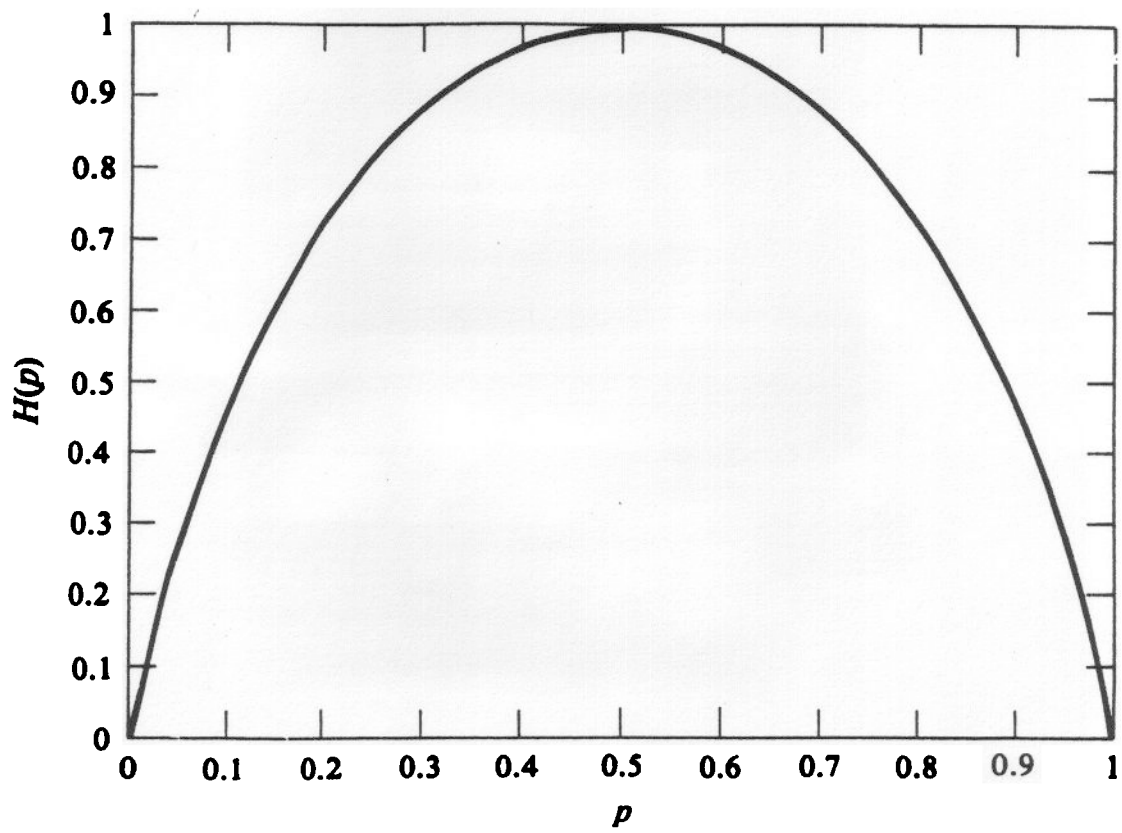


Figure 2.1. $H(p)$ versus p .

