Homework 2: Cryptanalysis

This homework is due **Wednesday, February 4** at **6 p.m.** and counts for 5% of your course grade. Late submissions will be penalized by 10% plus an additional 10% every 5 hours until received. You may submit one assignment late without this penalty. In either case, late work will not be accepted after 20 hours past the deadline. If you have a conflict due to travel, interviews, etc., please plan accordingly and turn in your homework early.

We encourage you to discuss the problems and your general approach with other students in the class. However, the answers you turn in must be your own original work, and you are bound by the Honor Code. Solutions should be submitted electronically via CTools in plain text format by completing the template at the end of this document.

Solve both of the following problems. You will probably want to write some short programs to help; submit them along with your answers. This might be a good opportunity to try Python, but you may use any common language or numerical package.

1. Here is some ciphertext that was produced with a Vigenère cipher:

```
GERGHDEVNIYMXQXUXGFPAXVVVOCBWGNUMDDXRVYBXOYRLWCGQIPNISRMQZNEWRYXVHEWCR
LRIISHCRPXFTIFMEWGOGRMRLTOAJIQTUJTPWNKZSHKSEXXSCCVNECMXRMFTSCJRGBGJHGS
GGBEUPLHNGRZNXMAZGFDRSPHCGIFEGHMSGASZXKVTGRSELSCAIBYRRKCVFTIWTQEAWNVXA
LNKVFTJEGXUHDTEEBFIHYWCXTHHGRVGWCGKEGBFBHCGHKZHNQYPARGSYXNVFBUGHRGKWPJ
MGRUOIYMAMVUGGXLTLHWCRGBTOIGSATERCMREXGISGEGBFBBMHRKEQGWTGHXFPNLLBEHTP
WRVKGIFIQBJQXNPVGVGDDQNMYSBYXVVJQDKTHMVFHAMRGTSPLHREVQIPMPTCSCEMAXVFXL
KZHUSGLGERGHDEVNIYMXQLRTMWAWFNLVRDLQNMYSBYXVVRZIFIBKPOCBGBFGIICVFVZSCA
ICKRQIGGRVIMERSTKRDWGGNEXCGGXUFJOGCHRLZUCCHNKFICBGBFGIIYXVHEOAFEEWESHQ
EFLLAERMBGJAPIMAZJIRFEYZFFXRLZLYOGBXBUISPIMAIIORRMPXSMPLCNWMSGQEERZHXQ
XUXFFTRMPTCZNNSFLZPACXBUISPIWHVYOHWWGXDPJRMGBJWCDINLZPACXBWFGDZCNGPYCM
AAIIORRMPTCATYRFVIMERSYHXMGCPNMVRICGUGFZDECUTJFPGWRWRBJKFRKFTACKNEZGHS
IFMYSTJIPMICCGGSKFBIGIEYFICBEGBFBLYWVGMCATIQBEORYWRBEHWCYABKSSQXNMVGLF
MPAHITQXVHESSULRMYSGPIDNZFXLKFNJDTAXRWTFXKMATCGIMTEHMWSCXUXZFSCGERGHXM
RXXPGIMPNPVBUMVPXDSCRMFNEQDLWGBKIIGSATCHWCISYRFVSIQMYOIRLVLZGPTMBERHXM
RBYKVTPMTAKCULSGUVWCEJBKTSSRSVGTFXKMATKSDLIFXCTPQKVOVBXLXUXWWURLNFVBSK
```

   Assume that encrypting with the key letter A results in no change, B results in an increment by one place in the alphabet, C results in an increment by two places, etc.

   What is the key? (Please show your work.)

2. Here is a table of the relative frequency of letters in English text:

A: 8.167%	B: 1.492%	C: 2.782%	D: 4.253%	E: 12.702%	F: 2.228%	G: 2.015%
H: 6.094%	I: 6.996%	J: 0.153%	K: 0.772%	L: 4.025%	M: 2.406%	N: 6.749%
O: 7.507%	P: 1.929%	Q: 0.095%	R: 5.987%	S: 6.327%	T: 9.056%	U: 2.758%
V: 0.978%	W: 2.360%	X: 0.150%	Y: 1.974%	Z: 0.074%		

Here is some plaintext:

```
ethicslawanduniversitypoliciestodefendasystemyouneedtobeabletothinklik
eanattackerandthatincludesunderstandingtechniquesthatcanbeusedtocompro
misesecurityhoweverusingthosetechniquesintherealworldmayviolatethelawo
rtheuniversitysrulesanditmaybeunethicalundersomecircumstancesevenprobi
ngforweaknessesmayresultinseverepenaltiesuptoandincludingexpulsioncivi
lfinesandjailtimeourpolicyineecsisthatyoumustrespecttheprivacyandprope
rtyrightsofothersatalltimesorelseyouwillfailthecourseactinglawfullyand
ethicallyisyourresponsibilitycarefullyreadthecomputerfraudandabuseactc
faaafederalstatutethatbroadlycriminalizescomputerintrusionthisisoneofs
everallawsthatgovernhackingunderstandwhatthelawprohibitsyoudontwanttoe
nduplikethisguyifindoubtwecanreferyoutoanattorneypleasereviewitsspolic
iesonresponsibleuseoftechnologyresourcesandcaenspolicydocumentsforguid
elinesconcerningproperuseofinformationtechnologyatumaswellastheenginee
ringhonorcodeasmembersoftheuniversitycommunityyouarerequiredtoabidebyt
```

The *population variance* of a finite population X of size N and mean μ is given by

$$\mathrm{Var}(\mathrm{X}) = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{x_i} - \mu)^2.$$

(a) What is the population variance of the relative letter frequencies in English text?

(b) What is the population variance of the relative letter frequencies in the given plaintext?

(c) For each of the following keys — yz, xyz, wxyz, vwxyz, uvwxyz — encrypt the plaintext with a Vigenère cipher and the given key, then calculate and report the population variance of the relative letter frequencies in the resulting ciphertext. Describe and briefly explain the trend in this sequence of variances.

(d) Viewing a Vigenère key of length k as a collection of k independent Caesar ciphers, calculate the mean of the frequency variances of the ciphertext for each one. (E.g., for key yz, calculate the frequency variance of the even numbered ciphertext characters and the frequency variance of the odd numbered ciphertext characters. Then take their mean.) Report the result for each key in part (c). Is the mean variance like those observed in part (b)? Part (c)? Briefly explain.

(e) Consider the ciphertext that was produced with key uvwxyz. In part (d), you calculated the mean of six variances for this key. Revisit that ciphertext, and calculate the mean of the frequency variances that arise if you had assumed that the key had length 2, 3, 4, and 5. Does this suggest a variant to the Kasiski attack? (Don't say no!) Briefly explain. □

Submission Template

```
# Problem 1:
key=XXXXXXXXXX

show_your_work_here ...

# Problem 2:
part_a_var_english=0.0000000
part_b_var_plaintext=0.0000000
part_c_var_ciphertexts=[0.0000000, 0.0000000, 0.0000000, 0.0000000, 0.0000000]
part_c_explain="briefly_describe_and_explain_trend ..."
part_d_means=[0.0000000, 0.0000000, 0.0000000, 0.0000000, 0.0000000]
part_d_explain="briefly_compare_and_explain_results ..."
part_e_means=[0.0000000, 0.0000000, 0.0000000, 0.0000000]
part_e_explain="briefly_explain_attack_variant ..."

show_your_work_here ...
```