This homework covers Lectures 1 and 2.

1. In this question you will practice your code-breaking skills.

   (a) The following ciphertext was derived by applying a substitution cipher.

   ```
   NTVW STVD NVJV MEE GW STV YJMNGWRJPPB STV QUVFSGPWF NTGIT VEGHMCVST
   TMY MEJVMYD MFKVY NVJV PO IPUJFV JVZVMSVY CD STV PSTVJF MWY STVD FPPW
   OPUWY STMS XMWV TMY WP GWSVEEGRVWIV SP RGAV STV FMWRUGWV TPZV PO RPPY
   TPNVAVJ NTGIT STV CVWVAPEVWIV PO TVJ TVMJS FURRVFSVY TMY WPS DVS
   YVFVJSVY TVJ FTV FSGEE VLZVISVY STMS GS NPUEY MEE VWY NVEE MWY STMS
   VAVJD BPJWGWR NPUEY CJGWR FPBV EVSSVJ VGSTVJ OJPB EDYGM PJ TVJ OMSTVJ
   SP VLZEMGW STVGJ ZJPIVVYGWRF MWY ZVJTMZF MWWPUWIV STV BMJJGMRV
   ```

   It is a paragraph from *Pride and Prejudice*, whose main character is `elizabeth`. Using this side information, decrypt the word `CVWVAPEVWIV`. (You may take advantage of the spaces.)

   (b) Decrypt the short pad ciphertext from page 7 of the Lecture 1 notes using the method described there. You may use the information that the key is 4 symbols long.

   (c) Complete the decryption of the short pad ciphertext from page 8 in the Lecture 1 notes.

2. These are the three most frequent digrams in ciphertexts obtained by encrypting the same 1715 character paragraph from *Pride and Prejudice* with (a) the shift cipher, (b) the substitution cipher, (c) the short pad with key length 3, and (d) the one-time pad. Which digram statistics arose from which cipher? Explain your reasoning.

1. digram	frequency	2. digram	frequency	3. digram	frequency	4. digram	frequency
TV	48	LX	22	LW	8	GD	48
VJ	43	HH	19	LH	8	DQ	43
ST	41	AY	18	VT	8	SG	41
GW	29	XA	18	BG	8	HM	29
PW	27	YV	17	NS	7	NM	27

3. Among the following four samplers, which produce indistinguishable outputs? Among those pairs that don't, what is the distinguisher with the highest advantage that you can come up with?

   Sampler A**:** Output x, where x is a pure random number modulo 6.
   Sampler B**:** Output $2x$ mod 6, where x is a pure random number modulo 3.
   Sampler C**:** Output $2x$ mod 6, where x is a pure random number modulo 6.
   Sampler D**:** Output x^2 mod 6, where x is a pure random number modulo 6.

4. This question is about message (in)distinguishability for the substitution cipher and the short pad cipher. In which of the following examples are encryptions of m_1 and m_2 indistinguishable? When they are not, what is the distinguisher of the highest advantage that you can think of?

   (a) $m_1 =$ `mama` and $m_2 =$ `papa` with the short pad cipher of length 4?

   (b) $m_1 =$ `mama` and $m_2 =$ `papa` with the substitution cipher?

   (c) $m_1 =$ `mamamama` and $m_2 =$ `mamapapa` with the short pad cipher of length 4?

   (d) $m_1 =$ `mamamama` and $m_2 =$ `mamapapa` with the substitution cipher?

   (e) $m_1 =$ `mamapapa` and $m_2 =$ `papamama` with the short pad cipher of length 4?