STAT 391
Final Exam
10:30 – 12:20 on June 8, 2006

©Marina Meilă
mmp@cs.washington.edu

Student Name:____

You are allowed 8 pages of notes.

Write your name clearly on of the notes pages; you will be asked to hand them in with the exam, but they will not be graded and will be returned to you in the statistics office after 4pm the day of the exam.

No electronic devices of any kind are allowed during the exam.

Any fact that was proved in the lectures or in the notes can be used without proof.

Prob.1 ____of 6

Prob.2 ____of 5

Prob.3 ____of 6

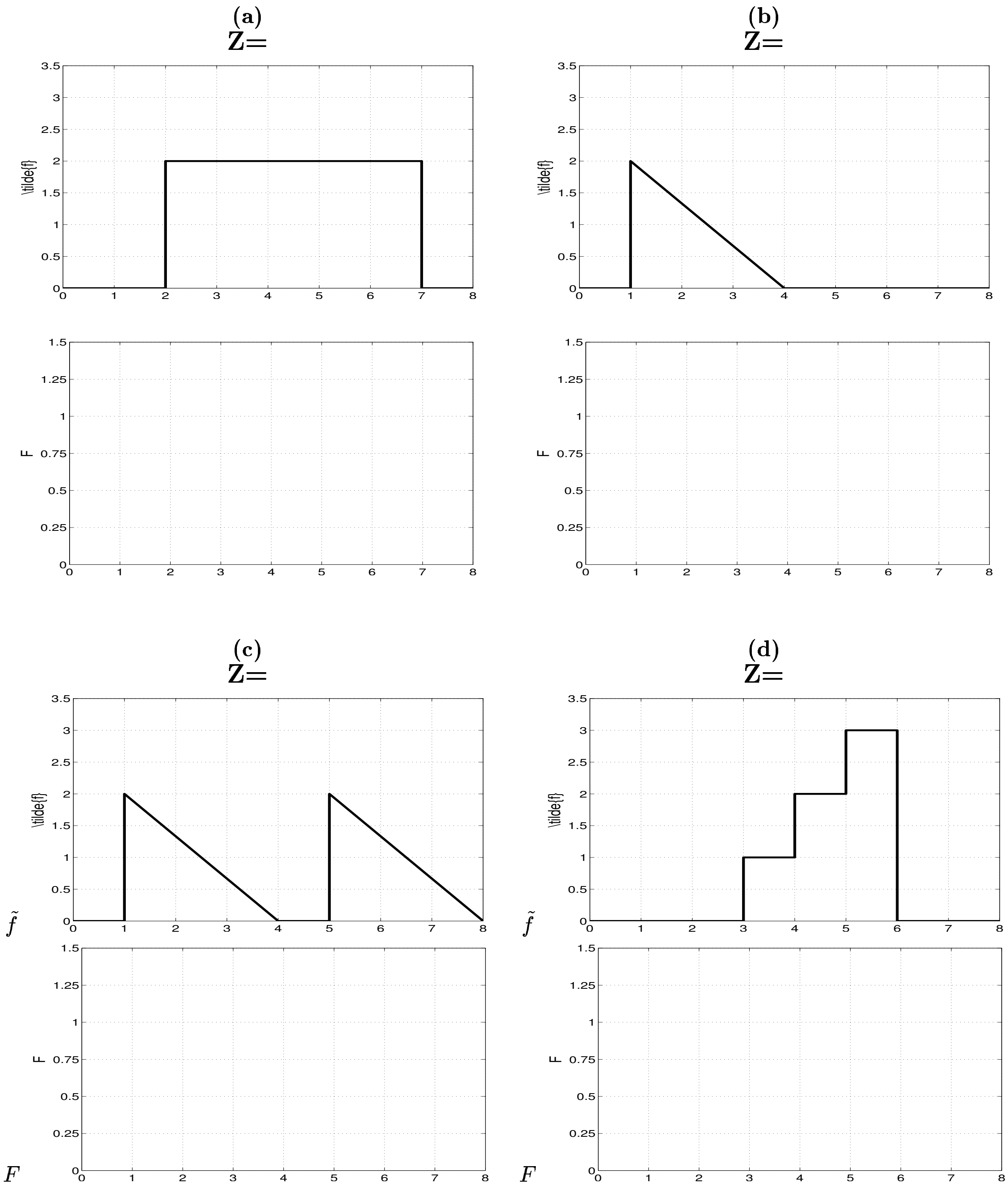
THIS IS UNFINISHED **Prob.4** ____of 5

Prob.5 ____of 4

Prob.6 ____of 9

Bonus 1 **TOTAL: ______of 36**

Problem 1 The following graphs represent unnormalized probability densities over the real axis. Assume that the densities are 0 outside the interval $[0, 8]$. Showing detail work is not required for this problem.



1.1 Compute the normalization constants for the four densities; write their values above the respective graphs. The normalization constant is the number Z so that $f = \frac{1}{Z}\tilde{f}$ is a probability distribution.

1.2 Draw the CDF's of the four densities in the corresponding plots.

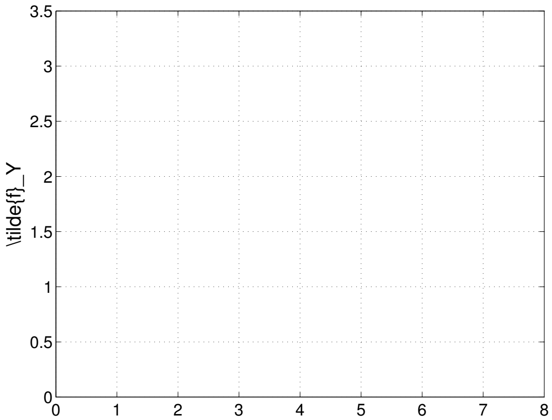
1.3 Compute the probability of the interval $[4, 6]$ under each of the four probability distributions.

1.4 Mark on graphs **a, d** the location of the median.

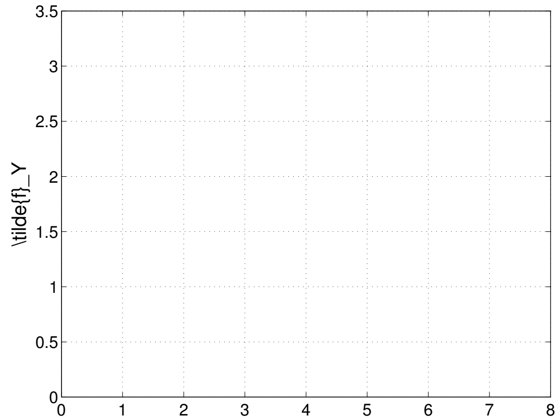
1.5 Mark on graphs **a, c** the location of the mean.

1.6 Let $Y = X - 1$ where X represents the original r.v in each of the four cases. Find S_Y and draw in the graphs below the (unnormalized) density of Y for each case.

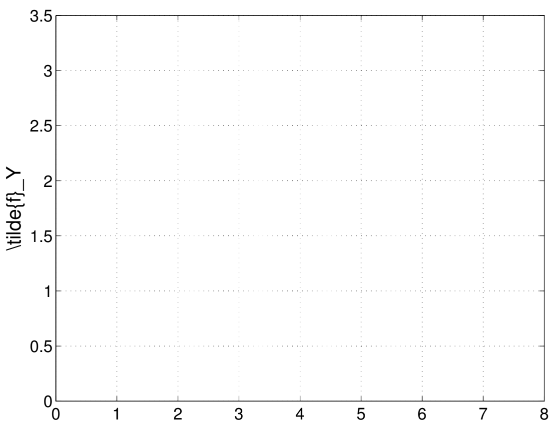
(a)
$\mathbf{S}_Y =$



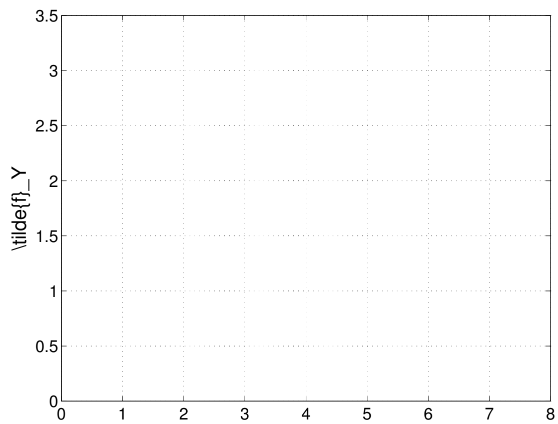
(b)
$\mathbf{S}_Y =$



(c)
$\mathbf{S}_Y =$

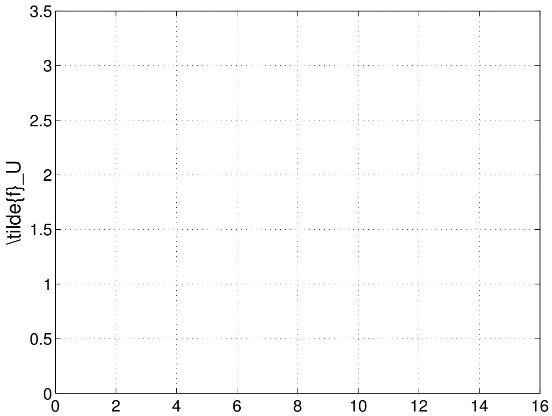


(d)
$\mathbf{S}_Y =$

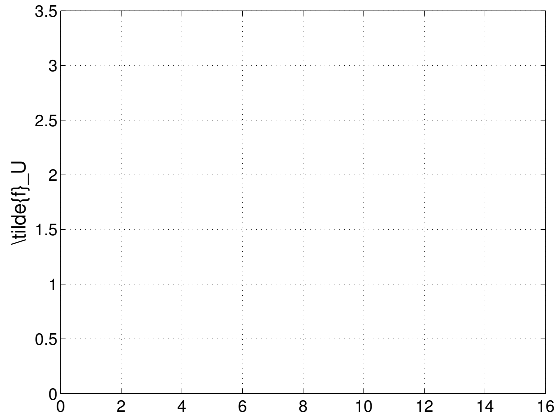


1.7 Let $U = 2X$ where X represents the original r.v in each of the four cases. Find S_U and draw in the graphs below the (unnormalized) density of U for each case.

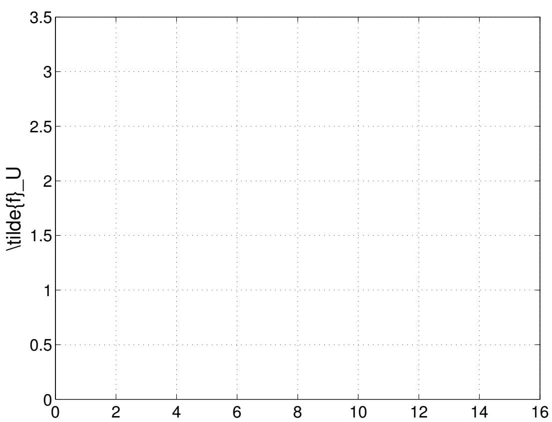
(a)
$\mathbf{S}_U =$
Z=



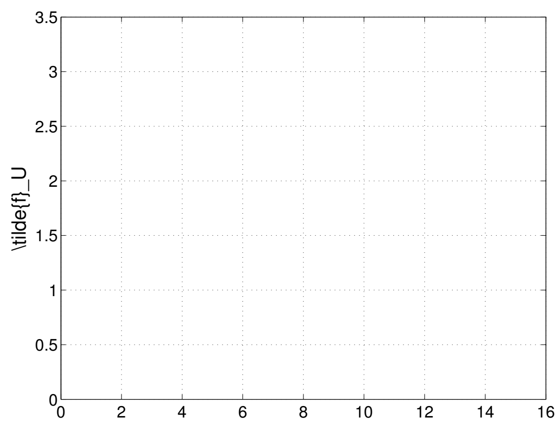
(b)
$\mathbf{S}_U =$
Z=



(c)
$\mathbf{S}_U =$
Z=



(d)
$\mathbf{S}_U =$
Z=

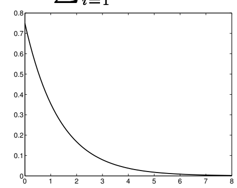


1.8 Compute the normalization constants for the four densities obtained in **1.7**; write their values above the respective graphs.

Problem 2 4 points

Compute and make neatly labeled graphs of the density estimate for each of the (very small) data sets given below. If you choose to draw an unnormalized f write the value of the normalization constant next to the plot. Show all your work.

Example Q: Find an exponential density estimate; data = {1, 1, 2}
Example A: $f_\lambda(x) = \lambda e^{-\lambda x}$ with $\lambda = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{3}{1+1+2} = \frac{3}{4}$



5.1 Estimate the parameters of the normal density estimate from the data = {2, 1, 1} and plot the resulting density. (A table with square roots of numbers 2–20 is available at the bottom of the page.) 2 p.

z	$\sqrt{z}$	z	$\sqrt{z}$	z	$\sqrt{z}$
2.0	1.41	8.5	2.92	14.5	3.81
2.5	1.58	9.0	3.00	15.0	3.87
3.0	1.73	9.5	3.08	15.5	3.94
3.5	1.87	10.0	3.16	16.0	4.00
4.0	2.00	10.5	3.24	16.5	4.06
4.5	2.12	11.0	3.32	17.0	4.12
5.0	2.24	11.5	3.39	17.5	4.18
5.5	2.35	12.0	3.46	18.0	4.24
6.0	2.45	12.5	3.54	18.5	4.30
6.5	2.55	13.0	3.61	19.0	4.36
7.0	2.65	13.5	3.67	19.5	4.42
7.5	2.74	14.0	3.74	20.0	4.47
8.0	2.83				

2.2 Find a kernel density estimate, for a square kernel with $h = 3$. Assume that the square kernel is $k(x) = \begin{cases} 1 & \text{if } x \in [-0.5, 0.5] \\ 0 & \text{otherwise} \end{cases}$ 2 p.

data = $\{-1, 1, 2\}$

Problem 3

3.1 Let A, B be events in a sample space S endowed with probability distribution P, and $0 < P(B) = q < 1$. Show that

$$A \perp B \quad \text{implies} \quad A \perp \overline{B}$$

In the above, $\overline{B} = S \setminus B$ and the symbol $\perp$ denotes probabilistic independence.

3.2 X, Y are random variables, $S_X = \{0,\, 1\}$, $S_Y = \{0,\, 1,\, 2\}$, $P_Y(y) > 0$ for all $y \in S_Y$. Show by a counterexample that

$$P_X(x) = P_{X|Y}(x|0) \text{ for all } x \in S_X \quad \not\Rightarrow \quad X \perp Y$$

Problem 4 Two continuous random variables X, Y taking values in $S_X = S_Y = [0, \infty)$ are described by the joint density

$$f_{XY}(x, y) = \frac{1}{Z} e^{-\lambda(x+y)} \quad x, y \in [0, \infty), \ \lambda > 0$$

Let $U = X + Y$. Answer the following questions, showing all your work.

For the notes: The univariate exponential distribution: $f(t) = \lambda e^{-\lambda t}$, $t \in [0, \infty)$, $\lambda \in (0, \infty)$. $E[t] = \frac{1}{\lambda}$, $Var\, t = \frac{1}{\lambda^2}$.

.1 Show that $X \perp Y$.

.2 Are X and U independent? Prove or disprove.

.3 Are X and U independent given Y? Prove or disprove.

.4 What are the marginal densities f_X, f_Y ?

.5 What is the expectation $E[U]$?

.6 What is the variance $Var\,U$?

.7 What is the probability density f_U of U?

Problem 5

Cryptographer Sofia is trying to open the entrance to the mysterious "treasure of Decebal"[1] The door has a lock that can be opened by a combination from the following letters:

l_1	l_2	l_3
A	A	E
B	I	M
M	M	P
	R	R
		T

That is, any combination (l_1, l_2, l_3) can be entered in the lock, but only one of them, "the key" word opens the door to the treasure.

Sofia knows the following:

The distribution of the first letter l_1 is uniform.

If the first letter is A, then the second letter l_2 is not a vowel, but it can be any of the consonants M,R with equal probability.

If the first letter is not A, the second letter is a vowel and $P(l_2 = A|l_1 \neq A) = \frac{1}{4}$, $P(l_2 = I|l_1 \neq A) = \frac{3}{4}$.

- The distribution of the third letter l_3 given the previous two is

$$P(\,l_3\,|\,l1\,<\,l_2\,) \;=\; \begin{cases} \text{T} & \text{w.p. } 0.6 \\ \text{E, M, P, R} & \text{w.p. } 0.1 \end{cases}$$

$$P(\,l_3\,|\,l1\,>\,l_2\,) \;=\; \begin{cases} \text{E} & \text{w.p. } 0.6 \\ \text{M, P, R, T} & \text{w.p. } 0.1 \end{cases}$$

5.1 What is the probability that the key starts with "MA"?

5.2 What is the probability of the word "BIT"?

[1] A king who died exactly 1900 years ago.

5.3 What is the probability of the event “B*T”, where * stands for any letter?

5.4 What is the probability that the key word ends in T?

5.5 What is the probability that the key word contains an A?

5.6 Sofia convinces the ghost who guards the treasure to help her. The ghost whispers the word in her ear, but she only can make the sound "..A.." and she's not even sure about that. But she is sure that the key word is one the following four MAT, MAP, BIT, ART.

The probability that Sofia hears "..A.." if the word contains an A is 0.9, and the probability of hearing "..A.." if the word doesn't contain A is 0.1.

Knowing this, can you help Sofia find the word that is most probably the key to the treasure?