STAT 391
Final Exam
2:30 – 4:20 on June 10, 2002
©Marina Meilä
mmp@cs.washington.edu

Student Name:______________________________

You are allowed 6 pages of notes. Calculators are not allowed.
Problem 1

Below are depicted four kernel estimates of the same density. The kernel type is the same in all plots.

Density estimate from \( D_1, h_1 \)

Density estimate from \( D_2, h_1 \)

Density estimate from \( D_1, h_2 \)

Density estimate from \( D_2, h_2 \)

a. What kernel was used to make the plots?

- [ ] Square  
- [ ] Epanechnikov  
- [ ] Gaussian

b. Which kernel width is larger?

- [ ] \( h_1 \)  
- [ ] \( h_2 \)

c. Which data set is larger?

- [ ] \( D_1 \)  
- [ ] \( D_2 \)
**Problem 2**

$X$ is a continuous random variable whose density $f(x)$ is proportional to the function $g(x)$ represented below, i.e.

$$f(x) = \frac{1}{Z} g(x)$$

Answer the following questions based on the plot. No proofs are necessary.

a. What is the normalization constant $Z$?

b. Compute the probability $P(X < 1)$
c. Compute the probability $P(1.2 \leq X \leq 2)$

e. Mark on the graph the position of $x_{1/2}$ the median of $X$. 
Problem 3

It is known that under $N(0,1)$ the normal distribution with mean 0 and covariance 1

\[
P(1, \infty) = p_1 \\
P(2, \infty) = p_2 \\
P(3, \infty) = p_3
\]

$Y$ and $Z$ are two random variables with

\[
Y \sim N(3,1) \\
Z \sim N(0,4)
\]

\[
\rho_{YZ} = \frac{1}{\sqrt{2}}
\]

[Hint: Make a drawing of $N(0,1)$, $f_Y$, $f_Z$ and think of probabilities as areas.]

\[\text{a. What is } P(Y < 1)?\]
b. What is $P(Y < 3)$?

c. What is $P(|Z| > 2)$?

d. What is $Cov(Y, Z)$?

e. Compute $P(Z < 0 | Y = 4)$. 
Problem 4

Professor Trelawney is teaching “Palmistry” and “Prediction with Tea leaves” at the Hogwarts School of Witchcraft and Wizardry. Denote by \( X \) the grade in Palmistry and by \( Y \) the grade in Tea leaves. The grades take values in the set \( \{ A, B, F \} \).

A new academic year is starting and Professor Trelawney needs to make some predictions to impress her class. Since her inner eye is a little tired, she wants to use previous year’s grade data for her predictions.

Below are the grades obtained by the \( n = 20 \) students in her previous class. They are sorted for her (and your) convenience.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>F</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>F</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the following use the Maximum Likelihood method if you need to estimate probabilities from data. Answer the other questions using the estimates obtained.

\[ \text{a. Estimate the distribution of } X, \text{ the grade in Palmistry, in last year’s class.} \]
b. Estimate the probability of the event “$X = Y$”.

c. Are the events “$X = Y$” and “$F$ in at least one test” independent? Justify your answer.

d. To pass each test, one needs a grade of $A$ or $B$ in that test. What is the probability that a student passes both tests?

e. What is the probability that a student passes Palmistry given that (s)he gets a $B$ in Tea Leaves?
f. Are $X$ and $Y$ independent? Justify your answer.


g. This year’s class has $n' = 40$ students. What is probability that all students pass both tests?

h. What is the expected number of students in this year’s class that pass both tests?
i. After this year’s test in Palmistry, while Professor Trelawney was taking a nap, Professor Snape sneaks in and by magic replaces some of her grades with grades from his own Potions test. Professor Snape assigns grades with the following probabilities:

\[
\begin{array}{c|ccc}
X & A & B & F \\
0.1 & 0.3 & 0.6 \\
\end{array}
\]

and the prior probability that any given grade was changed by him is 0.2.

Harry Potter gets a \textit{F} in Palmistry. What is the probability that his grade was magicked by Snape?
Problem 5

Paragon Pictures is planning to release two movies this summer season: “Attack From Planet Entropy” and “Rainier Raging”. Let $A$ be the event that “Attack” is a success. The probability of $A$ is

$$p_A = \frac{1}{10}$$

Let $R$ be the event that “Rainier” is a success. The probability of $R$ is

$$p_R = \frac{1}{5}$$

The summer season lasts 100 days and the movies’ planned release dates are $t_A = 20$ and $t_R = 60$. A movie is a success or a failure independently of any other movie from the same studio.

What isn’t known yet at Paragon Pictures is that the romantic comedy “Doll story” produced by a small independent studio is going to be a huge success this season, shadowing everything else after it is released. The release time $t$ of “Doll story” is unknown, and its density $f_t$ is uniform in the interval $[0, 100]$.

The probability that “Attack” is a success in the new conditions is

$$P(A|t) = \begin{cases} 
p_A \frac{t_A - t}{100} & \text{if } t \leq t_A \\
p_A & \text{if } t > t_A 
\end{cases}$$

“Rainier” is luckier because

$$P(R|t) = \begin{cases} 
p_R \frac{t_R - t}{60} & \text{if } t \leq t_R \\
p_R & \text{if } t > t_R 
\end{cases}$$

a. Compute the probability that “Attack” is a success.
(extra space; questions continue on the next page)
b. Make a neatly labeled plot of the probability density of $t$, the release time of “Doll Story” given that “Attack” was a success.
c. At time $t_A$ "Doll story" was not released yet. Could Paragon Pictures increase the probability of success of "Rainier raging" by changing its release date? More precisely, compute the time $t^*_R$ that maximizes the probability of success of "Rainier raging".