ASEN 5044 Statistical Estimation for Dynamical Systems
Fall 2018
Homework 4

Out: Thursday 09/27/2018 (posted on Canvas)

Due: Thursday 10/4/2018 (Canvas - **no credit for illegible submissions**)

Show all your work and explain your reasoning.

1. In the game of blackjack, the player is initially dealt two cards from a deck of ordinary playing cards. Without going into all the game's details, it is enough to know that that the best possible hand for a player to receive on the initial deal is a combination of an ace of any suite and any face card or 10. What is the probability that the player will be dealt this combination?

2. Discrete random variables X and Y can each take on integer values 1, 3, and 5. The joint probability table of X and Y is given below.

X	$Y=1$	$Y=3$	$Y=5$
1	1/18	1/18	1/18
3	1/18	1/18	1/6
5	1/18	1/6	1/3

a. Are the random variables X and Y independent?

b. Find the unconditional (marginal) probability $P(Y=5)$.

c. What is the conditional probability $P(Y=5|X=3)$?

3. Simon, Problem 2.3.

4. Simon, Problem 2.5.

5. Simon, Problem 2.15. Note: the `rand` function in Matlab is useful for generating uniformly distributed random numbers (see `help rand` or `doc rand` for more details – such uniform random number generators are commonly available in any programming environment). The `hist` command may also prove useful.

6. Simon problem 2.16.

Advanced Questions ***PhD students in the class MUST answer ALL questions below in addition to regular homework questions above*** – *non-PhD students are welcome to try any of these for extra credit (only given if all regular problems turned in on time as well). In either case, Submit your responses for these questions with rest of your homework, but make sure these are clearly labeled and start on separate pages –* ***indicate in the .pdf file name (per instructions posted on Canvas) and on the front page of your assignment if you answered these questions (as a PhD student, or for extra credit) so they can be spotted, graded and recorded more easily****.*

AQ1. Simon, Problem 2.7.

AQ2. Consider a continuous random variable x whose probability density function (pdf) $p(x)$ is given by a mixture model consisting of M weighted uniform pdfs,

$$p(x) = \sum_{i=1}^{M} w_i \cdot \mathcal{U}[a_i, b_i],$$

where $w_i \geq 0$, $\sum_{i=1}^{M} w_i = 1$, and $a_i < b_i$ for all mixture components ('mixands') $i = 1, ..., M$.

(a) Find general closed-form expressions for $\mathbb{E}[x]$ and $\text{var}[x]$ in terms of the means μ_i and variances σ_i^2 of the individual mixands $\mathcal{U}[a_i, b_i]$ (express in terms of w_i, a_i and b_i).

(b) Suppose the (w_i, a_i, b_i) parameters for $M = 7$ mixands are specified as follows, in order for components $i = 1, ..., 7$:

$$(0.1859, -4, -2),\ (0.0961, -2, -1),\ (0.1055, -0.75, 0),$$
$$(0.2104, 0, 1),\ (0.0678, 3, 5.5),\ (0.1950, 4, 6),\ (0.1393, -0.9, 7).$$

Plot the resulting values for $p(x)$ with these mixture parameters over the range $x \in [-6, 8]$, using an abscissa step size of $\Delta x = 0.001$ to evaluate at discrete grid points.

(c) Use (b)'s results to compute the following via numerical grid-based approximations:

- $\mathbb{E}[x]$ – how does this compare to the analytical result?
- $\text{var}(x)$ – how does this compare to the analytical result?
- $\mathcal{H}[p(x)] = \mathbb{E}[-\log p(x)]$ (the differential entropy; **hint:** think carefully about what should happen for $p(x) = 0$)
- $\text{KL}[p(x)||q(x)] = \mathbb{E}[\log \frac{p(x)}{q(x)}]$ (the Kullback-Leibler divergence between $p(x)$ and $q(x)$, which can be thought of as a kind of 'distance' measure between two pdfs, such that $\text{KL}[p(x)||q(x)] = 0$ iff $q(x) = p(x)$ and $\text{KL}[p(x)||q(x)] > 0$ for all $q(x) \neq p(x)$ with $q(x) > 0$ wherever $p(x) \geq 0$).

(d) Generate random sample sets of size $N = 100$, $N = 1,000$, and $N = 5,0000$ from $p(x)$, and plot the results on histograms with 1000 bins to compare to (b)'s results.

(e) Use each of the sample sets from (d) to produce direct Monte Carlo approximations of the quantities evaluated in (c). How do the results for each sample set size compare to each other and to the results from part (c)?