Homework # 4

ENAE 441

Instructions

Please answer the following questions. The following assignment has **TWO required Gradescope submissions**: `HW4` and `HW4-Analytic`.

Please submit your answers using the following naming conventions to Gradescope in two files:

1. `{your_name}.pdf` — derivations and / or written explanations of results
2. `submission.py` — the python script for the autograder (must be named `submission.py`!)

Note the type of question is indicated by its color:

- **Blue** questions are solved via pen and paper and uploaded in your **PDF file**.
- **Red** questions correspond to **autograded functions** that must be populated in `submission.py`.
- **Black** questions solutions are to be answered in `submission.py` but are **not autograded** (e.g. plots, discussions, etc)

Programming Questions

Each assignment has starter code available on ELMS. Your grade depends exclusively on the output of functions labeled `#REQUIRED`. Additional helper functions are included in the starter code to assist with debugging.

Questions (HW4-Analytic)

Answer the following questions and upload your work to the `HW4-Analytic` assignment on Gradescope.

Question 1. (15 pts) Consider two zero-mean uncorrelated random variables W and V with standard deviations σ_w and σ_v respectively. What is the standard deviation of the random variable $X = W + V$? Note: Do not assume gaussian distribution.

Question 2. (20 pts) Consider two scalar RVs X and Y.

a) Prove that if X and Y are independent, then their correlation coefficient $\rho = 0$.
b) Find an example of two RVs that are not independent but that have a correlation coefficient of zero.
c) Prove that if Y is a linear function of X then $\rho = \pm 1$.

Question 3. (25 pts) Consider the following function

$$f_{XY}(x, y) = \begin{cases} ae^{-2x}e^{-3y} & x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases}$$

and answer the following questions.[1]

a) Find the value of a so that $f_{XY}(x, y)$ is a valid joint probability density function.
b) Calculate $\bar{x}$ and $\bar{y}$. [2][3]
c) Calculate $\mathbb{E}[X^2], \mathbb{E}[Y^2]$ and $\mathbb{E}[XY]$.
d) Calculate the autocorrelation matrix of the random vector $Z = [X, Y]^T$
e) Calculate the variance σ_x^2, variance σ_y^2, and the covariance C_{XY}.
f) Calculate the autocovariance matrix of the random vector $Z = [X, Y]^T$.
g) Calculate the correlation coefficient between X and Y.

Question 4. (15 pts) Prove the following two results from lecture where x $\mathcal{N}(\bar{x}, \sigma_x^2)$ and e $\mathcal{N}(0, \sigma_e^2)$ and $y = cx + de$

a) $\text{cov}(X, Y) = \mathbb{E}[(x - \bar{x})(y - \bar{y})] = \mathbb{E}[XY] - \bar{x}\bar{y}$
b) $\text{var}(Y) = \mathbb{E}[(y - \bar{y})^2] = c^2\sigma_x^2 + d^2\sigma_e^2$

[1] Be sure stay organize and reuse information from earlier parts of this question... many intermediate expressions can be reused to reduce effort integrating.
[2] You'll need to use integration by parts
[3] Note: $\lim_{x\to\infty} cx\exp(-ax) = 0$ via L'Hopital where a, c are constants.

Questions (HW4-Programming)

Answer the following questions using Python and upload your work to the HW4 assignment on Gradescope as `submission.py`.

Question 5. (25 pts) Central Limit Theorem + Mappings

a) In python, generate a random variable x_1 distributed by a uniform distribution $x_i \sim \mathcal{U}[-1,1]$ using `np.random.uniform` function. Sample $N = 10$ points from this distribution, plot those points as a histogram.

b) Compute the *sample mean*, $\hat{\mu}$, and *sample variance*, $\hat{\sigma}^2$, of a sample set using functions `np.mean` and np.var functions, and determine if the reported values match the analytic mean and variance, $\mathbb{E}[x_1]$ and $\mathbb{E}[(x_1 - \mu_{x_1})^2]$ respectively? Repeat using $N = 10^i$ samples where $1 \leq i \leq 6$, reporting the sample mean. Report the values and explain what you observe.

c) Create three new random variables x_2, x_3, x_4 in addition to x_1, each also distributed from a uniform distribution $\mathcal{U}[-1,1]$. Sample $N = 100,000$ values from x_1 - x_4, and use these independent variables to compute a new set of random variables y_1, y_2, and y_3 defined as

$$y_1 = \frac{x_1 + x_2}{2} \tag{1}$$
$$y_2 = \frac{x_1 + x_2 + x_3}{3} \tag{2}$$
$$y_3 = \frac{x_1 + x_2 + x_3 + x_4}{4} \tag{3}$$
$$\tag{4}$$

Using the sampled values of x_1 through x_4, plot a histogram $p(x_1), p(y_1), p(y_2), p(y_3)$.

d) Explain what you see.

e) Transform y_3 into two new random variables z and q and plot the resulting distributions of $p(z)$ and $p(q)$ next to the original distribution $p(y_3)$ where

$$z = g(y) = 2y + 3 \tag{5}$$
$$q = f(y) = e^y \tag{6}$$

f) Explain how the distribution changes with both transformations, and if it makes sense.