MATH 180A: INTRO TO PROBABILITY
(FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Homework 8 Due TODAY

Final Exam: Monday, Dec 9, 11:30a–2:30p
in REC GYM

* Bring Student ID
* Seat Assignment on TritonEd (this weekend)
* 2 double-sided sheets of hand-written notes
* If possible, eat an early lunch.

Regrade requests (for final exam & HW8) NEXT QUARTER.

It is your responsibility to make sure all your grades are in order BEFORE NEXT FRIDAY.

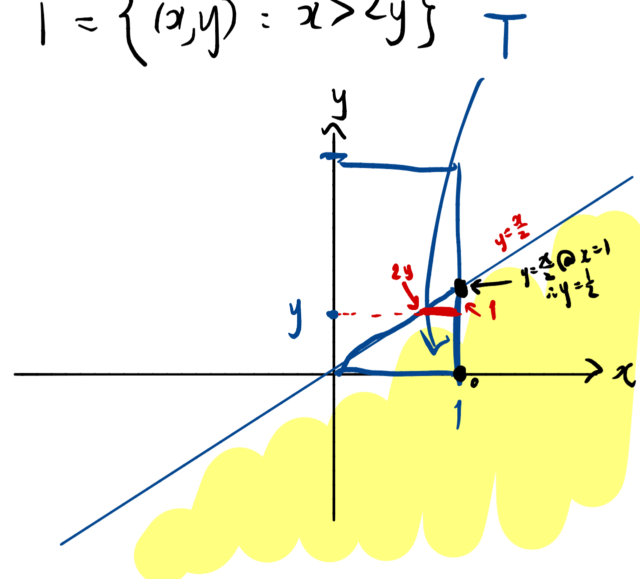
Please fill out your CAPEs!

1. Let X and Y be continuous random variables with joint density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2} + \frac{y}{4}, & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0, & \text{otherwise} \end{cases}$$

(a) (5 points) Compute $\mathbb{P}(X > 2Y)$.

$\{X>2Y\} = \{(X,Y)\in T\} \leftarrow \quad T = \{(x,y) : x > 2y\}$



$$\therefore \mathbb{P}(X>2Y) = \iint_T f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_0^{1/2} dy \int_{2y}^{1} dx \left(\frac{x}{2}+\frac{y}{4}\right)$$

$$= \int_0^{1/2} \left.\left(\frac{x^2}{4} + \frac{xy}{4}\right)\right|_{2y}^{1} dy$$

$$= \int_0^{1/2} \left(\frac{1-(2y)^2}{4} + \frac{(1-2y)y}{4}\right) dy$$

$$= \frac{1}{4}\int_0^{1/2} (1 - \underbrace{4y^2 + y - 2y^2}_{-6y^2})\,dy = \frac{1}{4}\left(y - 2y^3 + \frac{1}{2}y^2\right)\Big|_0^{1/2}$$

$$= \frac{1}{4}\left(\frac{1}{2} - 2\cdot\left(\frac{1}{2}\right)^3 + \frac{1}{2}\cdot\left(\frac{1}{2}\right)^2\right)$$

$$= \frac{1}{4}\left(\frac{1}{2} - \frac{1}{8}\right) = \frac{3}{32}.$$

Calculus

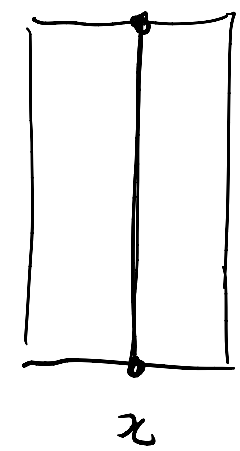
1. Let X and Y be continuous random variables with joint density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2} + \frac{y}{4}, & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0, & \text{otherwise} \end{cases}$$

(b) (5 points) What is the marginal density f_X of X?

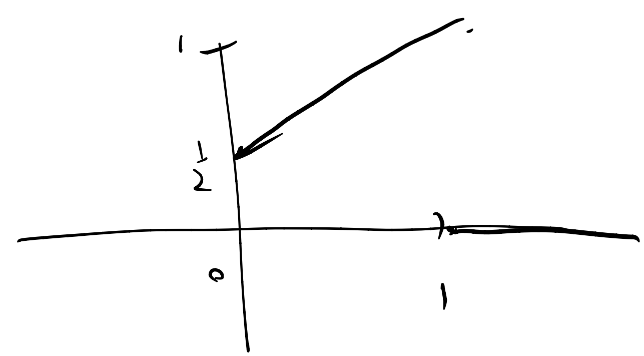
$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy \quad \text{if } 0 \le x \le 1;$$

0 otherwise.



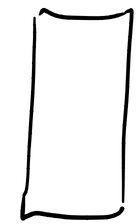
$$= \int_0^2 \left(\frac{x}{2} + \frac{y}{4}\right) dy$$

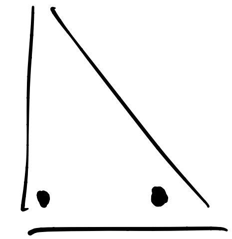
$$= \left(\frac{xy}{2} + \frac{y^2}{8}\right)\Bigg|_{0=y}^{2=y} = \frac{x \cdot 2}{2} + \frac{2^2}{8} = x + \frac{1}{2}.$$



1. Let X and Y be continuous random variables with joint density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2} + \frac{y}{4}, & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0, & \text{otherwise} \end{cases}$$





(c) (5 points) Are X and Y independent? Explain.

$$f_{(X,Y)}(x,y) = f_X(x)\, f_Y(y) \rightarrow f_Y(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{\frac{x}{2}+\frac{y}{4}}{x+\frac{1}{2}}$$

not a fn of y alone.

No. $\frac{x}{2}+\frac{y}{4}$ is not a product $u(x)v(y)$

S'pose it is. $\dfrac{\frac{x}{2}+\frac{0}{4}}{\frac{x}{2}+\frac{2}{4}} = \dfrac{\cancel{u(x)}\,v(0)}{\cancel{u(x)}\,v(2)}$ $\therefore$ $\dfrac{v(0)}{v(2)} = \dfrac{\frac{x}{2}}{\frac{x}{2}+\frac{1}{2}} = \dfrac{x}{x+1}$

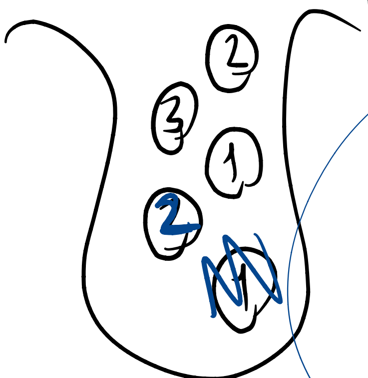
const.

not const.

$\lightning$

2. An urn contains 5 balls: two balls labeled 1, two balls labeled 2, and one ball labeled 3. You sample 2 balls from the bin, without replacement; the number on the first one is X, and the number on the second one is Y.

(a) (5 points) Compute the joint probability mass function of X and Y. (You may find it convenient to express it in the form of a chart.)



1st $\to$ X

2nd $\to$ Y

$P(X=1) = 2/5$

$P(X=2) = 2/5$

$P(X=3) = 1/5$

$$P(X=j, Y=k)$$
$$= P(Y=k \mid X=j)\,P(X=j)$$

$\{X=1\}$ $P(Y=1 \mid X=1) = 1/4$

$P(Y=2 \mid X=1) = 2/4 = 1/2$

$P(Y=3 \mid X=1) = 1/4$

$\{X=3\}$ $P(Y=3 \mid X=3) = 0$

$P((X,Y)=(j,k))$ $1 \le j,k \le 3$

	$X=1$	$X=2$	$X=3$
$Y=1$	$\frac{1}{4}\cdot\frac{2}{5} = 1/10$	$1/5$	$1/10$
$Y=2$	$\frac{1}{2}\cdot\frac{2}{5} = 1/5$	$1/10$	$1/10$
$Y=3$	$\frac{1}{4}\cdot\frac{2}{5} = 1/10$	$1/10$	0

2. An urn contains 5 balls: two balls labeled 1, two balls labeled 2, and one ball labeled 3. You sample 2 balls from the bin, without replacement; the number on the first one is X, and the number on the second one is Y.

(b) (5 points) Compute the probability mass function of X and the probability mass function of Y.

$p_X = p_Y$

	$X=1$	$X=2$	$X=3$	
$Y=1$	$1/4 \cdot 2/5 = 1/10$	$1/5$	$1/10$	$2/5$
$Y=2$	$\frac{1}{2} \cdot 2/5 = 1/5$	$1/10$	$1/10$	$2/5$
$Y=3$	$\frac{1}{4} \cdot 2/5 = 1/10$	$1/10$	0	$1/5$
$p_X \rightarrow$	$2/5$	$2/5$	$1/5$	$\uparrow p_Y$

2. An urn contains 5 balls: two balls labeled 1, two balls labeled 2, and one ball labeled 3. You sample 2 balls from the bin, without replacement; the number on the first one is X, and the number on the second one is Y.

(c) (5 points) Are X and Y independent? Justify your answer.

No!

$$\mathbb{P}(X=3, Y=3) = 0$$

But $\mathbb{P}(X=3)\,\mathbb{P}(Y=3) = \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$

$$\mathbb{E}(U_j) = 0 = \frac{2+(-2)}{2}$$

3. Let $U_1, U_2, \ldots, U_n, \ldots$ be independent, identically distributed random variables, each with the $\mathrm{Uniform}[-2,2]$ distribution. Let $S_n = U_1 + U_2 + \cdots + U_n$.

$$f_{U_1}(x) = \frac{1}{4}\mathbb{1}_{[-2,2]}(x)$$

(a) (5 points) Compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$.

$$\mathbb{E}(S_n) = \mathbb{E}(U_1 + U_2 + \cdots + U_n) = \mathbb{E}(U_1) + \mathbb{E}(U_2) + \cdots + \mathbb{E}(U_n) = 0$$

$$\mathrm{Var}(S_n) = \mathrm{Var}(U_1 + U_2 + \cdots + U_n)$$

$$\overset{\text{indepent}}{=} \mathrm{Var}(U_1) + \mathrm{Var}(U_2) \cdots + \mathrm{Var}(U_n) = \frac{4}{3}n.$$

$$\mathrm{Var}(U_1) = \mathbb{E}(U_1^2) - \cancel{\underset{0}{\mathbb{E}(U_1)^2}}$$

$$\int_{-2}^{2} x^2 \frac{1}{4}\, dx = \frac{1}{3}x^3\frac{1}{4}\Big|_{-2}^{2} = \frac{2}{4}\cdot\frac{2^3}{3} = \frac{16}{3\cdot 4} = \frac{4}{3}.$$

3. Let $U_1, U_2, \ldots, U_n, \ldots$ be independent, identically distributed random variables, each with the $\mathrm{Uniform}[-2,2]$ distribution. Let $S_n = U_1 + U_2 + \cdots + U_n$.

(b) (5 points) For any $\epsilon > 0$, what can you say about

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \epsilon\right)?$$

0

Chebyshev: $\mathbb{P}(|X - E(X)| \geq k\sigma(X)) \leq \frac{1}{k^2}$

$\mathbb{P}(|S_n - \underbrace{E(S_n)}_{=0}| \geq k\,\sigma(S_n)) \leq \frac{1}{k^2}$

$\sigma(S_n) = \sqrt{Var(S_n)} = \sqrt{\frac{4}{3}n}$

So $\mathbb{P}(|S_n| \geq k\sqrt{\frac{4}{3}n}) \leq \frac{1}{k^2}$.

$\therefore \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \underbrace{k\frac{\sqrt{\frac{4}{3}n}}{n^{2/3}}}_{\varepsilon}\right) \leq \frac{1}{k^2}$

$\therefore \frac{k\sqrt{\frac{4}{3}n}}{n^{2/3}} = \varepsilon$

$k = \frac{\varepsilon n^{2/3}}{\sqrt{\frac{4}{3}n}}$

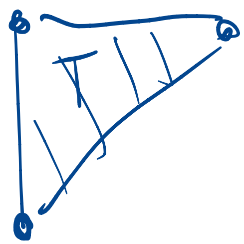
$\therefore \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \varepsilon\right) \leq \frac{1}{\left(\frac{\varepsilon n^{2/3}}{\sqrt{\frac{4}{3}n}}\right)^2}$

$= \left(\frac{\sqrt{\frac{4}{3}n}}{\varepsilon n^{2/3}}\right)^2 = \frac{\frac{4}{3}n}{\varepsilon^2 n^{4/3}} = \frac{\frac{4}{3}/\varepsilon^2}{n^{1/3}}$

$\therefore \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \varepsilon\right) \leq \frac{\text{const.}}{n^{1/3}} \to 0$ as $n \to \infty$.

4. Let T be the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(0,1)$, and $(1,1)$ (including the interior). Suppose that $P=(X,Y)$ is a point chosen uniformly at random inside of T.

(a) (5 points) What is the joint density function of (X,Y)? Use this to compute $\mathrm{Cov}(X,Y)$.



(X,Y) uniform in T

$$f_{(X,Y)}(x,y) = \frac{1}{\text{Area}(T)} \mathbb{1}_T(x,y)$$

$$= 2\,\mathbb{1}_T(x,y)$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{4} - \frac{1}{3}\cdot\frac{2}{3} = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}.$$

$$\mathbb{E}(XY) = \iint_T xy\, f_{(X,Y)}(x,y)\,dx\,dy = 2\iint_T xy\,dx\,dy = \frac{1}{4}$$

$$\mathbb{E}(X) = \iint_T x\cdot 2\,dx\,dy = \frac{1}{3} \qquad \mathbb{E}(Y) = \iint_T y\cdot 2\,dx\,dy = 2/3$$

4. Let T be the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(0,1)$, and $(1,1)$ (including the interior). Suppose that $P = (X,Y)$ is a point chosen uniformly at random inside of T.

(b) (5 points) Determine if X and Y are independent.

No! They are correlater:

$$\mathrm{Cov}(X,Y) = \frac{1}{36} \neq 0.$$

5. Suppose $X_1, X_2, \dots, X_n, \dots$ are i.i.d. random variables with mean $\mathbb{E}(X_j) = 0$ and variance $\mathrm{Var}(X_j) = 1$. Determine the following limits with precise justifications.

(a) (5 points) $\lim\limits_{n\to\infty} \mathbb{P}\left(-\frac{n}{4} \le X_1 + \cdots + X_n < \frac{n}{2}\right)$

5. Suppose $X_1, X_2, \ldots, X_n, \ldots$ are i.i.d. random variables with mean $\mathbb{E}(X_j) = 0$ and variance $\mathrm{Var}(X_j) = 1$. Determine the following limits with precise justifications.

(b) (5 points) $\lim_{n\to\infty} \mathbb{P}(X_1 + \cdots + X_n = 0)$

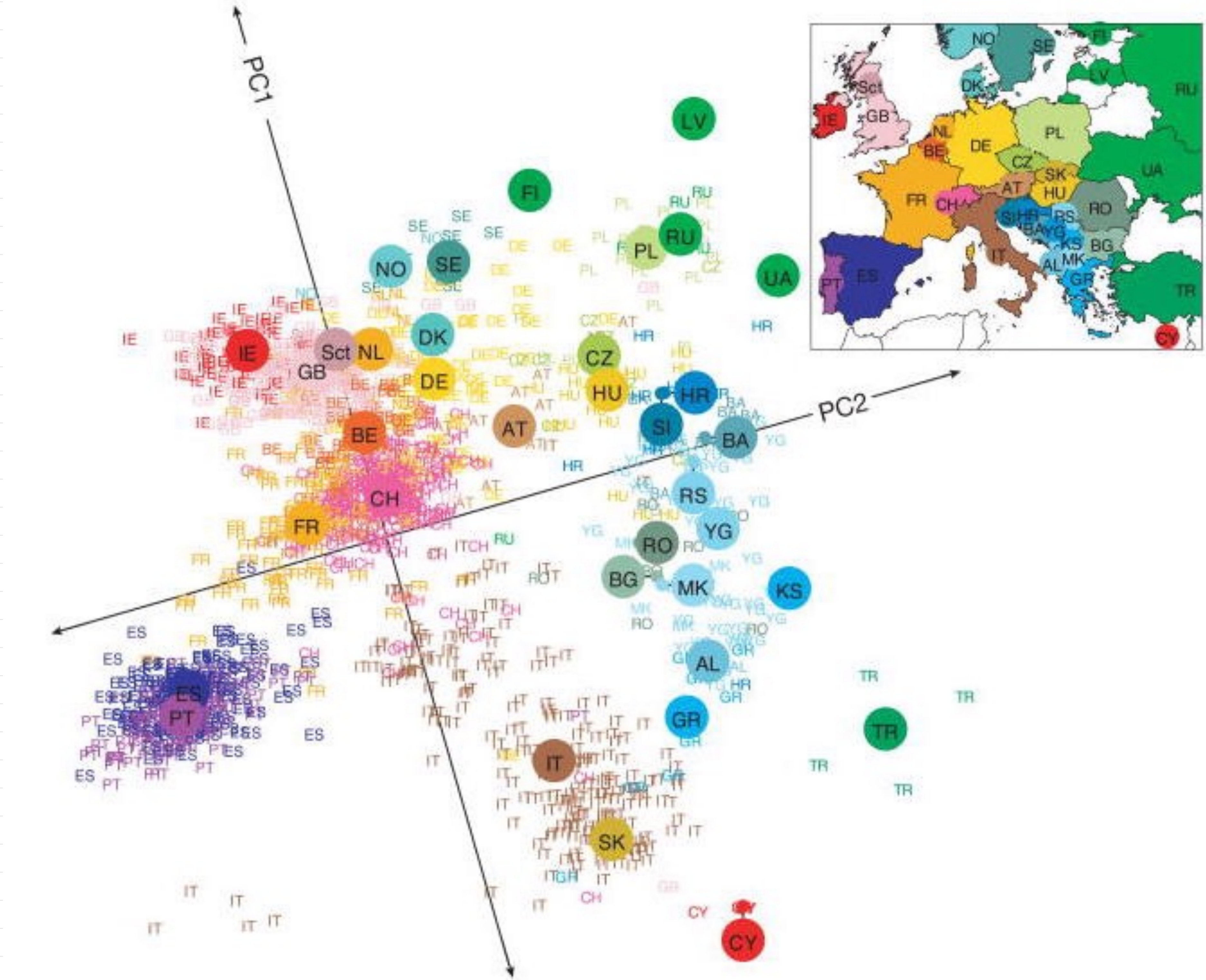
Please fill out your CAPEs !

And...

Sign up for my Winter 2020 *NEW COURSE*

MATH 182 / DSC 155

Hidden Data in Random Matrices



PC1
PC2
LV
FI
RU
PL
UA
NO
SE
DK
IE
Sct
NL
GB
DE
CZ
HU
HR
SI
BA
AT
BE
CH
FR
RS
YG
RO
BG
MK
KS
AL
GR
TR
IT
ES
PT
SK
CY