Information Retrieval, Assignment 1

Haris Jabbar, IMAT: 03647193
Juan Carlos Medina, IMAT: 03658400
Muhammad Zeeshan Karamat, IMAT: 03691742
Sukanya Raju, IMAT: 03692840

May 17, 2017

Problem 1

See python code and its comments.

Problem 2

For D = 2, the radius of the inner sphere is equal to the distance from the origin to a corner of the square minus 1/2, which is the radius of the sphere centered at the corner as shown in Figure 1. With the help of the Pythagorean theorem the distance from the center to a corner h is

$$h^2 = c^2 + c^2$$

Where $c = 1/2$, half of the size of the edge unit square and therefore

$$h = \frac{1}{\sqrt{2}}$$

The radius of the inner sphere that touches the four spheres is then

$$r_{center} = h - \frac{1}{2}$$

$$r_{center} = \frac{2 - \sqrt{2}}{2 * \sqrt{2}} = 0.207$$

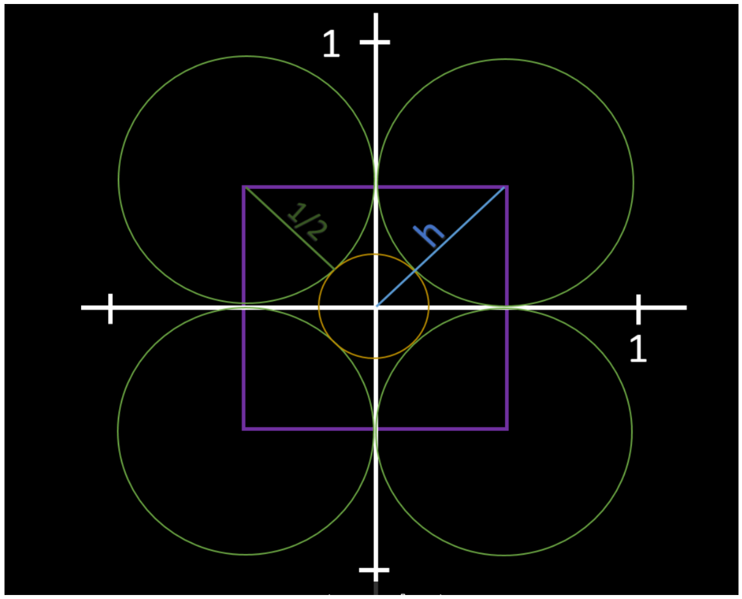

Figure 1: Unit square in 2D with spheres at its corners

For D = 3, the distance from the origin to the corner of the cube is greater than in two dimensions, in this case

$$h^2 = c^2 + c^2 + c^2$$

$$r_{center} = \frac{\sqrt{3}}{2} - \frac{1}{2} = 0.36$$

A general formula for the radius of the sphere that touches the spheres at the corners in n dimensions is

$$r_{center} = \frac{\sqrt{n}}{2} - \frac{1}{2}$$

For D = 9

$$r_{center} = \frac{\sqrt{9}}{2} - \frac{1}{2} = 1$$

This is counterintuitive, since on the main diagonals the sphere's radius can not be bigger than 1/2 . In higher dimensions this n-sphere would look as if it had spikes that go towards the corners of the hypercube. Remember that the volume of a unit sphere goes to zero as n goes to infinity and the normal rules that we are used to in 3 dimensions change when going to higher dimensions. From the general formula, we can see that when the dimension increases towards infinity, the radius also grows to infinity and the spheres around the corners of the hypercube are far away from the origin.

Problem 3

- **Part 1** Probability table

$$\sum_x \sum_y f(x,y)$$

= 0.4 + 0.14 + 0.05 + 0.02 + 0.26+ 0.13 = 1
As sum of all values of the table is 1, its a probability table.

- **Part 2**
Conditional Expectation

$$\mathbf{E}(y|x=2)[Y] = \sum_y Yp(y|x=2)$$

$$\frac{p(y=1,x=2)}{p(x=2)} = \frac{0.4}{0.59}$$

$$\frac{p(y=2,x=2)}{p(x=2)} = \frac{0.14}{0.59}$$

$$\frac{p(y=3,x=2)}{p(x=2)} = \frac{0.05}{0.59}$$

$$\sum_y Yp(y|x=2) = Y_1\frac{p(y=1,x=2)}{p(x=2)} + Y_2\frac{p(y=2,x=2)}{p(x=2)} + Y_3\frac{p(y=3,x=2)}{p(x=2)}$$

$$\sum_y Yp(y|x=2) = 1\frac{0.4}{0.59} + 2\frac{0.14}{0.59} + 3\frac{0.05}{0.59}$$

$$\sum_y Yp(y|x=2) = \frac{0.83}{0.59}$$

Probability :

$$P(x=1|y=3) = \frac{P(x=1,y=3)}{P(y=3)}$$

$$P(x=1|y=3) = \frac{0.13}{0.13+0.05}$$

$$P(x=1|y=3) = \frac{0.13}{0.18}$$

$$P(x=1|y=3) = 0.722$$

- **Part 3** For the given function to be joint density function, the two following properties need to statisfy :
1.

$$P(x,y) >= 0 \forall x,y$$

This is true from the equation
2.
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(x,y)dxdy = 1$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(x,y)dxdy = \int_{0}^{1/2}\int_{0}^{1} 1dxdy$$
$$= \int_{0}^{1/2} [x]_0^1 dy$$
$$= \int_{0}^{1/2} (1-0)dy$$
$$= \int_{0}^{1/2} dy$$
$$= [y]^1/2_0$$
$$= 1/2 - 0$$
$$= 1/2$$

Therefore the given function is not a joint distribution for two random variables.

- **Part 4** Marginal density function of y

$$\Pr\left(y\right) = \int_{0}^{y} 2e^{-(x+y)}dx$$
$$\Pr\left(y\right) = 2e^{-y}\left(\left|-e^{-x}\right|_0^y\right)$$
$$\Pr\left(y\right) = 2e^{-y}\left(1-e^{-y}\right)$$
$$\Pr\left(y\right) = 2e^{-y}\left(1-e^{-y}\right)$$

Marginal density function of x

$$\Pr\left(x\right) = \int_{x}^{\infty} 2e^{-(x+y)}dy$$
$$\Pr\left(x\right) = 2e^{-x}\left(\left|-e^{-y}\right|_x^\infty\right)$$
$$\Pr\left(x\right) = 2e^{-x}\left(e^{-x}-e^{\infty}\right)$$
$$\Pr\left(x\right) = 2e^{-x}\left(e^{-x}\right)$$
$$\Pr\left(x\right) = 2e^{-2x}$$

- **Part 5**

$$\Pr\big(x \le x_0 | y = y_0\big) = \frac{\int_{-\infty}^{x_0} f_{xy}(x,y)dx}{\int_{-\infty}^{\infty} f_{xy}(x,y)dx}$$

$$\Pr\big(x \le 2 | y = \frac{1}{2}\big) = \frac{\int_0^2 \frac{1}{15}(2x+2)dx}{\int_0^3 \frac{1}{15}(2x+2)dx}$$

$$\Pr\big(x \le 2 | y = \frac{1}{2}\big) = \frac{\int_0^2 (x+1)dx}{\int_0^3 (x+1)dx}$$

$$\Pr\big(x \le 2 | y = \frac{1}{2}\big) = \frac{|(\frac{x^2}{2}+x)|_0^2}{|(\frac{x^2}{2}+x)|_0^3}$$

$$\Pr\big(x \le 2 | y = \frac{1}{2}\big) = \frac{2+2-0}{\frac{9}{2}+3}$$

$$\Pr\big(x \le 2 | y = \frac{1}{2}\big) = \frac{8}{15}$$