DATA 1010
Problem Set 6
Due 19 October 2018 at 11 PM

Problem 1

In this problem, we will justify the idea that the mean of a random variable is the best constant estimator for the random variable, as measured by average squared error.

Suppose that X is a random variable for which $\mathbb{E}[X^2] < \infty$. Show that the minimum of the function $f(a) = \mathbb{E}[(X-a)^2]$ occurs at the value $a = \mathbb{E}[X]$.

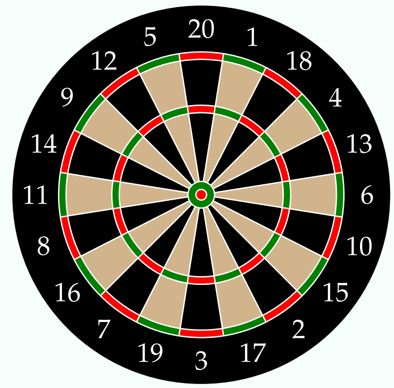
Problem 2

Suppose that the probability density function for the random point where your dart hits the dartboard* $D = \mathbb{R}^2$ is given by

$$f(x,y) = \frac{1}{\pi}\mathrm{e}^{-x^2-y^2},$$

where the origin is situated at the dartboard's bull's eye, and where x and y are measured in inches (this function is positive everywhere in $\mathbb{R}^2$, so the "dartboard" includes the disk shown as well as the (infinite) wall it is mounted on—this is realistic insofar as one can indeed hit the wall with a dart throw). Find the probability of scoring triple 20 on your next throw.

Note: the triple 20 region is the smaller of the two thin red strips in the sector labeled "20". The inner and outer radii of this thin strip are 3.85 inches and 4.2 inches, respectively.



Problem 3

Find the expected distance from the origin to a point (X,Y) selected uniformly at random from the triangle T with vertices at $(0,0)$, $(1,0)$, and $(1,1)$. For fun, make your best estimate of the answer before you do any calculation, and at the end you can comment on whether your estimate ending up being low or high.

Hints: recall that $\mathbb{E}[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ if (X,Y) has a joint distribution specified by the probability density function f. Also, you probably want to set the integral up in polar coordinates, and feel free to use a symbolic computation engine (such as Wolfram Alpha) to perform the actual integration.

Problem 4

Suppose that X and Y are random variables whose joint distribution is given by the density $f(x,y) = \frac{3}{2}(x^2+y^2)$ on the unit square $[0,1]^2$.

(a) Find the probability density function of the distribution of X.

(b) Find the probability of the event that $X \geq \frac{1}{2}$ and $Y \geq \frac{1}{2}$.

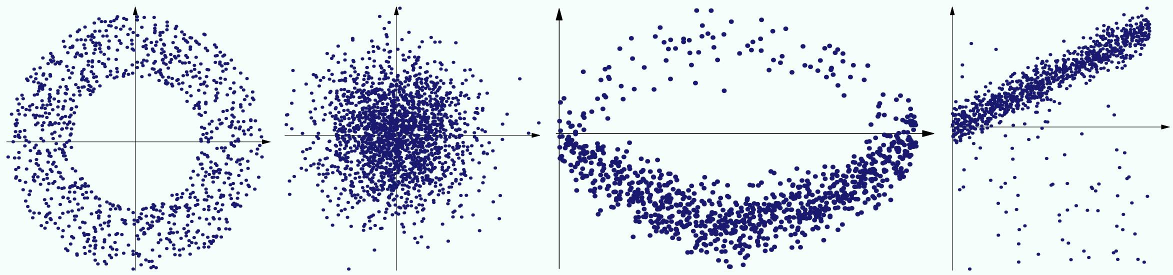
Problem 5

Consider two events A and B. Show that the indicator random variables $\mathbf{1}_A$ and $\mathbf{1}_B$ are positively correlated if $\mathbb{P}(A \mid B) > \mathbb{P}(A)$ and are negatively correlated if $\mathbb{P}(A \mid B) < \mathbb{P}(A)$.

Problem 6

Suppose that many independent samples are drawn from the joint distribution of two random variables X and Y, and the results are as shown in the first figure below. Sketch, by shading, your best guess of the density function of the distribution of (X,Y). Also sketch a graph of the function $x \mapsto \mathbb{E}[Y \mid X = x]$.

Repeat the exercise for each of the remaining figures.



Problem 7

Suppose that X and Y have joint PDF $f(x, y) = \frac{3}{2}y$ on the upper unit disk (that is, the set of points which have positive y-coordinate and are less than one unit from the origin).

(a) Verify that f is indeed a probability density function.

(b) Find the density of the distribution of X.

(c) Find the conditional density of Y given $X = x$.

(d) Find $\mathbb{E}[Y \mid X]$.

Problem 8

You're tasked with loading a pile of stones into your truck. No one has looked carefully at the pile yet, but based on past experience you decide to model the number of stones as a random variable N with $\mathbb{E}[N] = 50$. The stone weights are also random variables (independent of N), and each one has a mean of 10 pounds. Let X be the total weight of the stones in the pile.

(a) For any positive integer $n \geq 1$, find $\mathbb{E}[X \mid N = n]$. Use the result to find $\mathbb{E}[X \mid N]$ in terms of the random variable N.

(b) Use the fact that $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid N]]$ to find $\mathbb{E}[X]$.