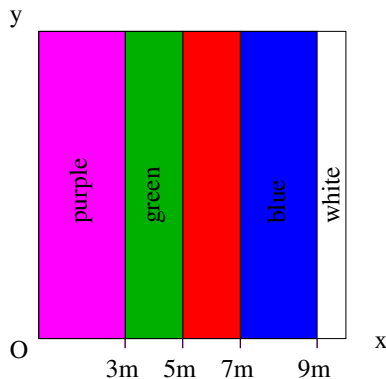


RANDOM VARIABLES

Consider an experiment where a vanishingly small object lands at random in a square region with an area of 100cm^2 . Call the sample space Ω and assign random variables as follows:

- X: *half* the distance from where the object lands to the y -axis in cm
- Y: the distance from where the object lands to the x -axis in cm
- C: the cosine of the distance from where the object lands to the x -axis in cm divided by 1cm (that is, $\cos(Y/1\text{cm})$)
- R: the distance from where the object lands to the origin in cm
- L: 0 or 1 if the object lands at a point whose y -coordinate is $\leq 5m$. Note that this random variable *depends* on another, Y
- S: divide the square into five rectangular regions as in the picture, with S giving the indicated *color* of the region in which the object lands.



Think about the definition of each of the random variables. They measure things about *actual results* from an experiment based on chance: the object lands and then you have at it with a ruler. Mathematically, they are *maps* from the sample space to a space of possible measurements for that random variable.

- (1) For each random variable, identify its codomain (the set of values it can take) write it as a map; for example, $X:\Omega \rightarrow \{0, 2\text{cm}\}$

To think about the sample space more abstractly, just imagine it as a bag of distinguishable points. As a bag of points, it's difficult figure out which point is which: even though they're distinguishable, you'd like to have a way to *organize* the points. "Aha!" you cry, I can assign to each point a pair of numbers that will let me arrange them in a square. These assignments are just functions: maps from a set to the

real numbers. In particular, you've given *coordinate functions* for the sample space.

But remember, *continuous random variables* are maps from the sample space to the real numbers too; is there a difference? It's clear that the joint random variable (X, Y) can be thought of as coordinates on the sample space, but (X, C) cannot. This is because the functions (X, C) cannot distinguish between the points that we describe in terms of the 'usual' coordinate functions as $(0, 0)$ and $(0, 2\pi cm)$.

With your new and complete grasp on random variables, answer the following.

- (2) Write down the density functions for X and Y , and write down precisely what it means for the object to land in the region *at random*. Notice that you use coordinate functions to describe "at random"!
- (3) Compute the density function for R . For the remainder of this question assume that R is uniformly distributed and explain what changes we would observe over the course of many iterations of the experiment. Write a program to simulate this experiment 1000 times and plot the results *in (x, y) coordinates* (note that you will need to define another random variable, θ – assign it a uniform density as well).
- (4) Are X and Y independent? Are X and R ?
- (5) What is the *event* described by $L = 1$? Notice that it is a collection of outcomes: the random variable is a map

$$L: \Omega \rightarrow \{0, 1\},$$

and we want the *event* $E \subset \Omega$ that is defined to be $E := \{\omega \in \Omega \mid L(\omega) = 1\}$. Are L and Y independent?

- (6) Compute the distribution function for S . Are S and X independent? Are S and Y ? Try to explain why in words for both cases.