Continuous Joint Distributions

Of course joint variables don’t have to be discrete only, they can also be continuous. As an example: consider throwing darts at a dart board. Because a dart board is two dimensional, it is natural to think about the X location of the dart and the Y location of the dart as two random variables that are varying together (aka they are joint). However since x and y positions are continuous we are going to need new language to think about the likelihood of different places a dart could land. Just like in the non-joint case continuous is a little tricky because it isn’t easy to think about the probability that a dart lands at a location defined to infinite precision. What is the probability that a dart lands at exactly \((X=456.234231234122355, Y = 532.12344123456)\)?

Let’s build some intuition by first starting with discritized grids. On the left of the image above you could imagine where your dart lands is one of 25 different cells in a grid. We could reason about the probabilities now! But we have lost all nuance about how likelihood is changing within a given cell. If we make our cells smaller and smaller we eventually will get a second derivative of probability: once again a probability density function. If we integrate under this joint-density function in both the x and y dimension we will get the probability that x takes on the values in the integrated range and y takes on the values in the integrated range!

Random variables \(X\) and \(Y\) are Jointly Continuous if there exists a Probability Density Function (PDF) \(f_{X,Y}\) such that:

\[
P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x,y) \, dy \, dx
\]

Using the PDF we can compute marginal probability densities:

\[
f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y) \, dy
\]

\[
f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b) \, dx
\]

Example

Let’s make a weight matrix used for Gaussian blur. In the weight matrix, each location in the weight matrix will be given a weight based on the probability density of the area covered by that grid square in a 2D Gaussian with variance \(\sigma^2\). For this example let’s blur using \(\sigma = 3\).
In image processing, a Gaussian blur is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise.

Gaussian blurring with StDev = 3, is based on a joint probability distribution:

**Joint PDF**

\[
f_{X,Y}(x,y) = \frac{1}{2\pi \cdot 3^2} e^{-\frac{x^2+y^2}{2 \cdot 3^2}}
\]

**Joint CDF**

\[
F_{X,Y}(x,y) = \Phi\left(\frac{x}{3}\right) \cdot \Phi\left(\frac{y}{3}\right)
\]

Each pixel is given a weight equal to the probability that \(X\) and \(Y\) are both within the pixel bounds. The center pixel covers the area where \(-0.5 \leq x \leq 0.5\) and \(-0.5 \leq y \leq 0.5\). What is the weight of the center pixel?

\[
P(-0.5 < X < 0.5, -0.5 < Y < 0.5) \\
= P(X < 0.5, Y < 0.5) - P(X < 0.5, Y < -0.5) \\
- P(X < -0.5, Y < 0.5) + P(X < -0.5, Y < -0.5) \\
= \Phi\left(\frac{0.5}{3}\right) \cdot \Phi\left(\frac{0.5}{3}\right) - 2\Phi\left(\frac{0.5}{3}\right) \cdot \Phi\left(\frac{-0.5}{3}\right) \\
+ \Phi\left(\frac{-0.5}{3}\right) \cdot \Phi\left(\frac{-0.5}{3}\right) \\
= 0.5662^2 - 2 \cdot 0.5662 \cdot 0.4338 + 0.4338^2 = 0.206
\]