Math 126 Problem 7: Deriving the Normal Distribution

Spring 2027

Submission Instructions

Please upload your work to Gradescope by **11:59pm on Monday, February 15, 2027 (Pacific time).**

- You are encouraged to discuss ideas and work with your classmates. However, you *must acknowledge* your collaborators at the top of each solution on which you collaborated with others and you *must write* your solutions independently.
- Your solutions to theory questions must be written legibly, or typeset in LaTeX or markdown.
- I recommend that you write your solutions in a Python notebook using Google Colab.
- You should submit your solutions as a *single PDF* via the assignment on Gradescope.

Grading: The point of the problem set is for *you* to learn. To this end, I hope to disincentivize the use of LLMs by *not* grading your work for correctness. Instead, you will grade your own work by comparing it to my solutions. This self-grade is due the Friday *after* the problem set is due, also on Gradescope.
Original notebook: Google Colab.

Deriving the Normal Distribution



In this problem, we will derive the density function of the normal distribution from the example of a dartboard. Let $f : \mathbb{R}^d \to [0, 1]$ be a probability density function that describes the probability a dart lands at the point (x, y). We want our probability density function f to have two properties:

- **Radial Symmetry:** The probability that a dart lands at a point depends only on the distance between the point and the origin.
- **Independence:** Coordinates are independent e.g., knowing the x-coordinate does not give us information about the y-coordinate.

From the radial symmetry property, we can conclude that

$$f(x,y) = f(r) = f(\sqrt{x^2+y^2})$$

where $r = \sqrt{x^2+y^2}$ is the distance between the origin and (x,y).

From the independence property, we can conclude that $f(x,y) = g(x)h(y)$ for some functions g and h. Further, the radial symmetry property tells us that $f(x,y) = f(y,x) = g(y)h(x)$ so g and h must be the same function. That is,

$$f(x,y) = g(x)g(y).$$

Part A

Up to rescaling, we can assume that $g(0) = 1$. Use this fact and plug in the point $(r,0)$ to conclude that f and g are the same function.

Part B

Define the function $h(x) = f(\sqrt{x})$. In particular,

$$f(\sqrt{x^2+y^2}) = f(x)f(y) \Leftrightarrow h(x^2+y^2) = h(x^2)h(y^2).$$

Use this property to show that

$$h(x_1 + x_2 + \ldots + x_n) = h(x_1)h(x_2) \cdot \ldots \cdot h(x_n).$$

Then prove that $h(n) = h(1)^n$.

Without loss of generality, let $h(1) = b$.

Part C

Let p and q be any integers. Prove that

$$h\left(\frac{p}{q}\right) = b^{\frac{p}{q}}.$$

Hint: Consider the expression

$$h\left(\frac{p}{q} \cdot q\right).$$

As long as h is continuous, conclude that $h(x) = b^x$ for any real number since the rational numbers are dense in the real number line.

Part D

You have shown that $h(x) = b^x$. Further, $h(x) = e^{cx}$ for some real number c. Then $f(x) = h(x^2) = e^{cx^2}$. Recall this implies the density function of x and y is $f(x, y) = f(x)f(y) = e^{c(x^2+y^2)}$.

For f to be a valid probability density function, what constraint do we have on c? If we wanted to make the distribution more concentrated, how should we change c?

Except for the normalization, we have explained every part of the multivariate normal distribution given below

$$f(x, y) = \frac{1}{\sigma^2 \cdot 2\pi} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

Why is there a normalization?