1 Continuous Joint Densities

The joint probability density function of two random variables X and Y is given by $f(x,y) = Cxy$ for $0 \le x \le 1, 0 \le y \le 2$, and 0 otherwise (for a constant C).

(a) Find the constant C that ensures that $f(x,y)$ is indeed a probability density function.

(b) Find $f_X(x)$, the marginal distribution of X

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

(d) Are X and Y independent?

2 Sum of Independent Gaussians

In this question, we will introduce an important property of the Gaussian distribution: the sum of independent Gaussians is also a Gaussian.

Let X and Y be independent standard Gaussian random variables. Recall that the density of the standard Gaussian is

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\Bigl(-\frac{x^2}{2}\Bigr).$$

(a) What is the joint density of X and Y?

(b) Observe that the joint density of X and Y, $f_{X,Y}(x,y)$, only depends on the quantity $x^2 + y^2$, which is the distance from the origin. In other words, the Gaussian is *rotationally symmetric*. Next, we will try to find the density of $X+Y$. To do this, draw a picture of the Cartesian plane and draw the region $x+y \le c$, where c is a real number of your choice.

(c) Now, rotate your picture clockwise by $\pi/4$ so that the line $X+Y=c$ is now vertical. Redraw your figure. Let X' and Y' denote the random variables which correspond to the $\pi/4$ clockwise rotation of (X,Y) and express the new shaded region in terms of X' and Y'.

(d) By rotational symmetry of the Gaussian, (X',Y') has the same distribution as (X,Y). Argue that $X+Y$ has the same distribution as $\sqrt{2}Z$, where Z is a standard Gaussian. This proves the following important fact: *the sum of independent Gaussians is also a Gaussian*. Notice that $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ and $X+Y \sim \mathcal{N}(0,2)$. *In general, if X and Y are independent Gaussians, then $X+Y$ is a Gaussian with mean $\mu_X+\mu_Y$ and variance $\sigma_X^2+\sigma_Y^2$.*

(e) Recall the CLT:

> If $\{X_i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 < \infty$, then:
>
> $$\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}} \xrightarrow{\text{in distribution}} \mathcal{N}(0,1) \qquad \text{as } n\to\infty.$$

Prove that the CLT holds for the special case when the X_i are i.i.d. $\mathcal{N}(0,1)$.

3 Binomial Concentration

Here, we will prove that the binomial distribution is *concentrated* about its mean as the number of trials tends to ∞. Suppose we have i.i.d. trials, each with a probability of success $1/2$. Let S_n be the number of successes in the first n trials (n is a positive integer), and define

$$Z_n := \frac{S_n - n/2}{\sqrt{n}/2}.$$

(a) What are the mean and variance of Z_n?

(b) What is the distribution of Z_n as $n\to\infty$?

(c) Use the bound $\mathbb{P}[Z>z] \le (\sqrt{2\pi}z)^{-1}\,\mathrm{e}^{-z^2/2}$ when Z is normally distributed in order to bound $\mathbb{P}[S_n/n > 1/2+\delta]$, where $\delta > 0$.