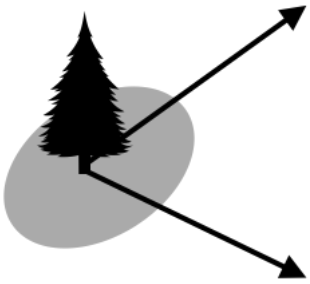
TD 2

September 9, 2024

Exercise 1

We want to model positions of fallen tree leaves on the ground.



Firstly, we assume that leaves are distributed uniformly inside a disk B_1 of radius 1 m. We denote (Z_1, Z_2) the 2-dimensional random variable corresponding to their coordinates.

1 **Generate $N = 1000$ samples of (X, Y) uniformly in $[-1, 1] \times [-1, 1]$ and plot them.**

2 **Use (X, Y) to generate $N = 100$ samples of (Z_1, Z_2) and plot them.**

3 **What is $\mathbb{P}\big((X, Y) \in B_1\big)$? How can we use that to get an approximated value of π ?**

4 **Plot a horizontal line of height π and its approximation with n from 10 to 500.**

Consider $(X_1, Y_1), \ldots, (X_n, Y_n)$ independent variables following the same distribution as (X, Y) and $C_i = 1$ if $(X_i, Y_i) \in B_1$ and $C_i = 0$ otherwise. $\overline{C_n} = \frac{1}{n} \sum_{i=1}^{n} C_i$.

5 What is the distribution of C_i ? Distribution of $\sum_{i=1}^{n} C_i$?

6 Compute and plot these discrete probabilities for $n = 10$, $n = 100$ and $n = 1000$. (Use scipy.special.binom **for binomial coefficients and** barplot **for plots.)**

7 What is $\mathbb{P}(a \leq \overline{C_n} \leq b)$?

8 What would be an exact symmetric confidence interval for $\mathbb{E}[C_i]$ of level $1-\alpha$? Define a Python function that computes it depending on α and n.

9 From which Gaussian distribution these barplots get closer while n gets big? Add the corresponding Gaussian curve to each of your previous barplots.

10 What would be an asymptotic symmetric confidence interval ? Define a Python function that computes it and compare results.

Now assume leaves are more likely to fall close to the tree in such a way that $(Z_1, Z_2) \sim \mathcal{N}\big((0,0), I_2\big)$

11 What is the distribution of (Z_1^2, Z_2^2) and $Z_1^2 + Z_2^2$?

12 Deduce $\mathbb{P}\big((Z_1, Z_2) \in B_1\big)$.

13 How can we approximate π using samples from Z_1 or Z_2 random variables ?

Finally assume that $\sigma_1^2 = \mathrm{Var}(\mathrm{Z}_1)$ and $\sigma_2^2 = \mathrm{Var}(\mathrm{Z}_2)$ are unknown (while being still independent and centered).

14 Write a Python function that returns estimators of these parameters with associated confidence intervals ?