Statistics, Fall 2025 - Practice Session 1
Sampling & Sufficiency

TA: Stylianos Daskalinas

Email: stylianosdaskalinas@gmail.com

Office hours by appointment. Office D3, 3rd floor , Bld. B

25/9/2025

Problem 1

Suppose we toss a fair coin n times.

1. What is the probability of getting Heads at every throw for n times.
2. Suppose that the coin is faulty, and the probability of getting Heads is only 25%. How many throws do we need to get Heads at least once with probability of at least 90%?

Problem 2

Are the following distributions members of the exponential family?

- Gamma distribution

- Cauchy distribution

- U-Quadratic distribution

- Rayleigh distribution

Provide mathematical reasoning for your answers.

Problem 3

Let $\{X_i\}_{i=1}^n$ be a sequence of IID random variables.

1. If $X_i \sim G(\lambda)$, where the distribution G has the PDF $f(x;\lambda) = \lambda x^{\lambda-1}\mathbb{1}(x)_{(0,1)}$, come up with a sufficient statistic for λ.
2. If $X_i \sim \Gamma$, with both parameters unknown. Find two jointly sufficient statistics for a and β.

Problem 4

Let $\{X_i\}_{i=1}^N$ be an IID sequence of Bernoulli random variables with parameter p and let N be Poisson-distributed random variable with parameter λ. Calculate the expectation of

$$\sum_{i=1}^{N} X_i$$

Problem 5

Suppose we have some random variable X, with a distribution belonging in a subset of the exponential family, with density

$$f(x;a) = g(a)e^{ax}$$

Show that

$$\mathbb{E}[X] = -\frac{d}{da}\log(g(a))$$

Notes

- The symbol $\mathbb{1}_A(x)$ is used for the indicator function over a set A, i.e.

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

- For the sufficient statistics exercises, remember that the goal is to come up with functions of the data points that do not include unknown parameters, and that we need as many statistics as the number of unknown parameters in the joint density.

- For a pair of dependent random variables X and Y, and an (integrable) function $Z = \psi(X, Y)$, the Law of Iterated Expectations states

$$\mathbb{E}[Z] = \mathbb{E}_X[\mathbb{E}_Y[Z|X]] = \mathbb{E}_Y[\mathbb{E}_X[Z|Y]]$$

- The Cauchy distribution has density

$$\frac{1}{\pi}\frac{\gamma}{\gamma^2 + (x-a)^2}$$

- The U-Quadratic distribution has density

$$a(x-b)^2\mathbb{1}_{[a,b]}(x)$$

- The Rayleigh distribution has density

$$\frac{x}{a}\exp(-\frac{x^2}{2a})\mathbb{1}_{[0,\infty)}(x)$$