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%?

Solution

Each toss is an IID Bernoulli trial, i.e. for each toss, we have

$$X_i = \begin{cases} \text{Heads} & \text{with probability } p \\ \text{Tails} & \text{with probability } 1-p \end{cases}$$

with $p = 0.5$. Then, we have

$$\Pr((X_1 = \text{Heads}) \cap (X_2 = \text{Heads}) \cap \cdots \cap (X_n = \text{Heads})) = \prod_{i=1}^{n} \Pr(X_i = \text{Heads}) = p^n$$

since each toss is independent from the previous one.

For the second question, since we only care about getting Heads at least once and not in which throw, we can use the Binomial distribution, i.e.

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

where X is the random variable of the number of times we get Heads in n tosses. We're looking for an n such that

$$0.9 \leq \Pr(X \geq 1)$$

First, we have that

$$\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - \Pr(X = 0) = 1 - \binom{n}{0} p^0 (1-p)^{n-0} = 1 - (1-p)^n$$

Solving for n we have

$$0.9 \leq 1 - (1-p)^n \Rightarrow (1-p)^n \leq 0.1 \Rightarrow n \log(1-p) \leq \log(0.1) \Rightarrow n \geq \frac{\ln(0.1)}{\ln(1-p)}$$

Substituting the faulty coin probability we get

$$n \geq \frac{\ln(0.1)}{\ln(1-0.25)} \approx 8.0039$$

Therefore, we need at least $n = 9$ tosses to get Heads at least once with probability 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.

Solution

For this exercise, all we need is the general density specification of the exponential family and some factorization. A random variable's distribution belongs in the exponential family if its density has the general specification

$$f(x;\theta) = a(x)b(\theta) \exp\left(\sum_{i=1}^{k} f_k(x) g_k(\theta)\right)$$

Then consider the following manipulations.

Gamma distribution: Yes, since

$$\begin{aligned}\frac{\beta^a}{\Gamma(a)}x^{a-1}e^{-\beta x} &= \frac{\beta^a}{\Gamma(a)}\exp\Big((a-1)\log(x)-\beta x\Big)\\ &= \Big(1\Big)\Big(\frac{\beta^a}{\Gamma(a)}\Big)\exp\Big((a-1)(\log(x))+(-\beta)(x)\Big)\end{aligned}$$

Cauchy distribution:
No, since the density of a Cauchy random variable is

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

The value within the logarithm cannot be decomposed in order to separate the parameters and the value of the variable, therefore the Cauchy distribution does not belong in the exponential family.

U-Quadratic distribution: No, since

$$a(x-b)^2 = a\exp(2\log(x-b))$$

Again, we cannot break the contents of the logarithm down into separate items, therefore this distribution is also not in the exponential family.

Rayleigh distribution: Yes, since

$$\frac{x}{a}e^{-\frac{x^2}{2a}} = \Big(x\Big)\Big(\frac{1}{a}\Big)\exp\Big((-\frac{1}{2a})(x^2)\Big)$$

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.

Solution

First, recall the Fisher-Neyman Factorization theorem. A statistic $T = \phi(X_1, \dots X_n)$ is sufficient if the joint PDF of the random sample can be written as

$$f(x_1, \dots, x_n; \theta) = h(x_1, \dots, x_n) g(\phi(x_1, \dots, x_n), \theta)$$

For the first case, consider the joint PDF for the random sample (assuming that all $x_i \in (0,1)$)

$$f(x_1, \dots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i, \lambda) = \prod_{i=1}^{n} \lambda x_i^{\lambda-1} = \lambda^n \Big(\prod_{i=1}^{n} x_i\Big)^{\lambda-1} = \left(\lambda^n \Big(\prod_{i=1}^{n} x_i\Big)^{\lambda}\right)\left(\Big(\prod_{i=1}^{n} x_i\Big)^{-1}\right)$$

Since the right parentheses term does not depend on λ, the Factorization theorem holds and the product

$$\prod_{i=1}^{n} x_i$$

is a sufficient statistic for λ.

For the second case, consider the PDF of the Gamma distribution

$$f(x; a) = \frac{\beta^a}{\Gamma(a)} x^{a-1} e^{-\beta x} \mathbb{1}(x)_{[0,\infty)}$$

The joint PDF for the random sample (ignoring the indicator) will be

$$f(x_1, \dots, x_n; a) = \prod_{i=1}^{n} f(x_i; a) = \prod_{i=1}^{n} \frac{\beta^a}{\Gamma(a)} x_i^{a-1} e^{-\beta x_i} = \Big(\frac{\beta^a}{\Gamma(a)}\Big)^n \prod_{i=1}^{n} x_i^{a-1} \exp(-\beta \sum_{i=1}^{n} x_i)$$

The jointly sufficient statistics become apparent once we untangle the observation values and parameters in the joint density

$$f(x_1, \dots, x_n; a) = \Big(\frac{\beta^a}{\Gamma(a)}\Big)^n \prod_{i=1}^{n} x_i^{a-1} \exp(-\beta \sum_{i=1}^{n} x_i) =$$
$$\Big(\frac{\beta^a}{\Gamma(a)}\Big)^n \exp\Big((a-1)\sum_{i=1}^{n} \ln(x_i) - \beta \sum_{i=1}^{n} x_i\Big) =$$
$$\left\{\Big(\frac{\beta^a}{\Gamma(a)}\Big)^n \exp\Big((a-1)\sum_{i=1}^{n} \ln(x_i) - \beta \sum_{i=1}^{n} x_i\Big)\right\}\left\{1\right\}$$

The obvious statistics are

$$T_1 = \sum_{i=1}^{n} \ln(x_i) \text{ and } T_2 = \sum_{i=1}^{n} x_i$$

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

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

Solution

For this problem, recall the Law of Total Expectation (very useful in both Statistics and Econometrics)

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

In our case, we have two categories of random variables, N and X_i. We can apply the Law of Total Expectation as follows:

$$\mathbb{E}\Big[\sum_{i=1}^{N} X_i\Big] = \mathbb{E}_N\Big[\mathbb{E}_X\Big[\sum_{i=1}^{N} X_i\Big|N\Big]\Big] = \mathbb{E}_N\Big[\sum_{i=1}^{N} \mathbb{E}_X[X_i|N]\Big]$$
$$= \mathbb{E}_N\Big[\sum_{i=1}^{N}[1 \times p + 0 \times (1-p)]\Big] = \mathbb{E}_N[Np] = p\mathbb{E}_N[N]$$

Since N is Poisson-distributed with parameter λ, we can easily calculate the expected value

$$\mathbb{E}_N[N] = \sum_{n=0}^{\infty} ne^{-\lambda}\lambda^n \frac{1}{n!} = \mathbb{E}_N[N] = \sum_{n=0}^{\infty} ne^{-\lambda}\lambda^{n-1+1} \frac{1}{n!}$$
$$= \lambda e^{-\lambda} \sum_{n=0}^{\infty} n\lambda^{n-1} \frac{1}{n!} = \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{1}{n!}\Big(\frac{d}{d\lambda}\lambda^n\Big) =$$
$$\frac{1}{\lambda} e^{-\lambda} \frac{d}{d\lambda}\Big(\sum_{n=0}^{\infty} \frac{1}{n!}\lambda^n\Big) = \lambda e^{-\lambda} \frac{d}{d\lambda}\Big(e^{\lambda}\Big) = \lambda$$

Thus, returning to the problem, the expectation of the sum is

$$\mathbb{E}\Big[\sum_{i=1}^{N} X_i\Big] = p\lambda$$

Problem 5

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

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

where $g(a)$ is a positive differentiable function. Show that

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

Solution

Start from $\int_{-\infty}^{\infty} f(x; a)dx = 1$. take the first derivative with respect to a

$$\begin{aligned}
&\frac{d}{da}\int_{-\infty}^{\infty} f(x; a)dx = \frac{d}{da}1 \Rightarrow \int_{-\infty}^{\infty}[g'(a)e^{ax} + xg(a)e^{ax}]dx = 0 \Rightarrow \\
&\int_{-\infty}^{\infty} g'(a)\frac{g(a)}{g(a)}e^{ax}dx + \int_{-\infty}^{\infty} xg(a)e^{ax}dx = 0 \Rightarrow \\
&\frac{g'(a)}{g(a)}\int_{-\infty}^{\infty} g(a)e^{ax}dx + \int_{-\infty}^{\infty} xg(a)e^{ax}dx = 0 \Rightarrow \\
&\frac{g'(a)}{g(a)}1 + \mathbb{E}[X] = 0 \Rightarrow \mathbb{E}[X] = -\frac{d}{da}\log(g(a))
\end{aligned}$$

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

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

- The Rayleigh distribution has density

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