1 Probability and Likelihood [3]

1.1 Definition

- Probabilities defined over discrete sets of events: $\mathrm{Prob}(dice = 1|dice \in \{1, 2, 3, 4, 5, 6\}) = 1/6$.
- Probabilities with respect to continuous variables: $\mathrm{Prob}(X = 5.0|X \in [1.0, 6.0]) = 1/\infty = 0$.
  - Probability density function (PDF): $y = f(x)$.
    $\mathrm{Prob}(x \in [2.0, 4.0]|f(x) = 0.2, \text{ for } 1.0 \le x \le 6.0) = \int_{2.0}^{4.0} f(x)dx = 0.2 \times 2 = 0.4$, where $f(x)$ is shown in Fig. 1(a).

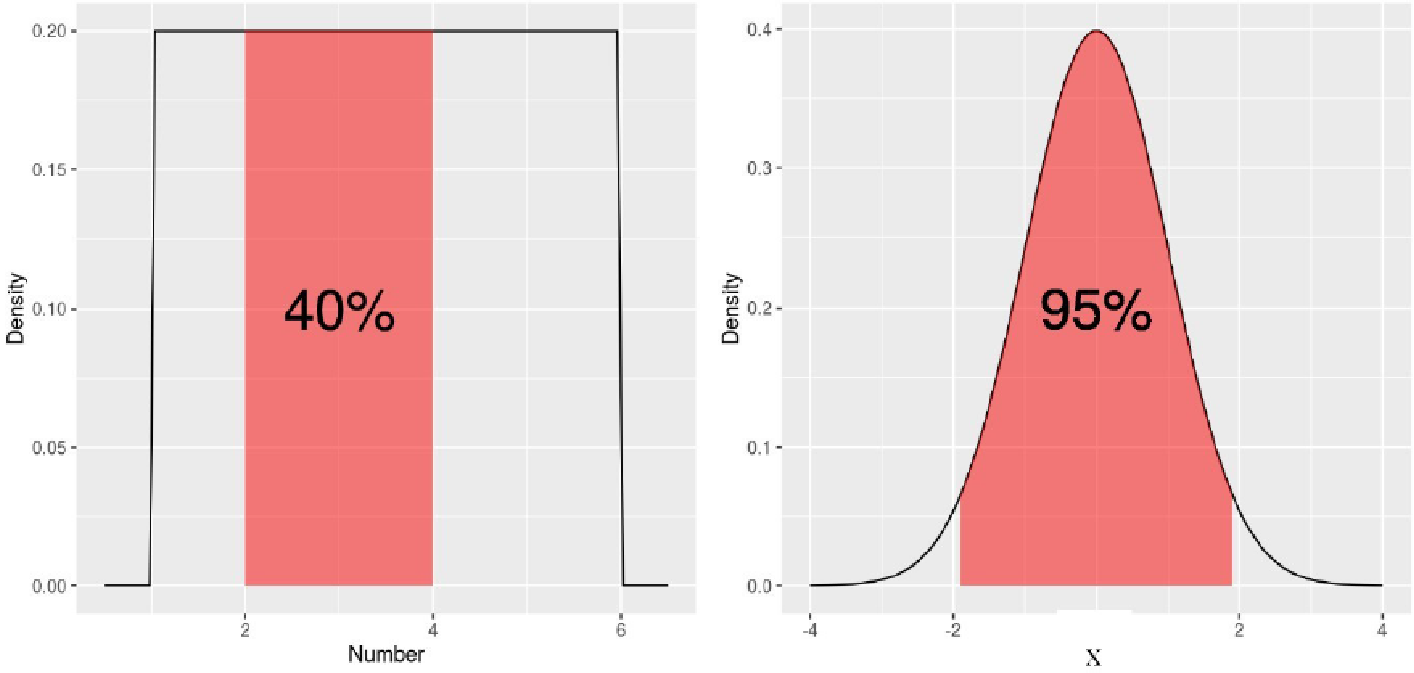

Figure 1: Probability density functions (PDFs) of continuous variables: (a) uniform distribution in $[1.0, 6.0]$, $y = 0.2$, for $1.0 \le x \le 6.0$ and (b) standard normal distribution, $y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

1.2 Maximum Likelihood Estimation

Given an experiment (event or system), estimate the parameter (or a set of parameters) that maximizes the likelihood of the experiment.

Example 1. Estimation of the probability of an uneven coin using multiple trials
Let p be an unknown probability of head when tossing an uneven coin, then the probability of tail is equal to $1-p$. To estimate p, that is considered as a parameter of the coin tossing system, we had the head 400 times and tails 600 times out of 1000 trials. The estimation problem is illustrated in Fig. 2.

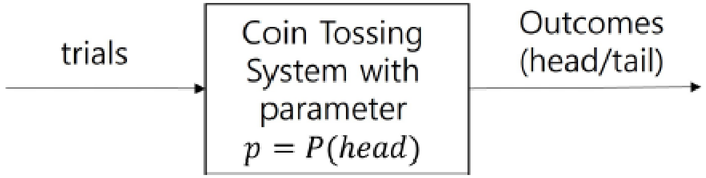

Figure 2: Estimation of uneven coin probability

The estimation process includes the following steps:

1. Get outcomes by trials. (400 times head, 600 times tail)
2. Define the likelihood as a function of p, such as$L = {}_{1000}C_{400}p^{400}(1-p)^{600}$ as shown in Fig. 3.

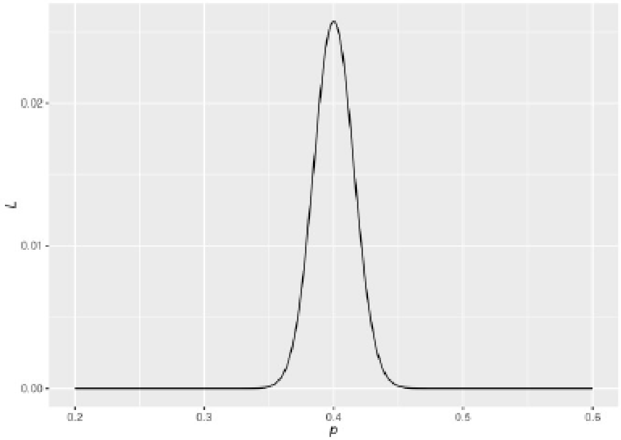

Figure 3: Likelihood function, $L = {}_{1000}C_{400}p^{400}(1-p)^{600}$

3. To maximize L, $\frac{dL}{dp} = 0$ yields the solution. ($p = 0.4$)

The estimated parameter maximizes the likelihood of given outcomes. In a continuous case, the likelihood is equivalent to the probability density function of the outcome, which states that $p = 0.4$ maximizes the probability of the outcome (400 head, 600 tail). $\square$