CGS698C, Assignment 01

Himanshu Yadav

Part 1: Probability

1.1 In an experiment, two coins are tossed simultaneously.

(A coin toss gives you either heads (denoted by H) or tails (denoted by T).)

(a) What is the sample space of the two-coin-tossing experiment?

(b) What is the event space of the experiment? List all the possible events.

(c) Suppose that all the elementary events in the sample space have equal probabilities.
(For example, if there are three possible outcomes in the sample space Ω such that $\Omega = \{x_1, x_2, x_3\}$, then $P(\{x_1\}) = P(\{x_2\}) = P(\{x_3\})$)

i. What is the probability of occurrence of each elementary event?

ii. What is the probability of the event that at least one head appears?

iii. What is the probability of the event that exactly one head appears?

Part 2: Discrete random variables

2.1 In a visual word recognition experiment, a participant has to recognize whether the word shown on screen is a meaningful word (e.g., "book") or a non-word (e.g., "bktr"). The participant is asked to answer "yes" if the shown string is a meaningful word, and "no" if it is a meaningless non-word. Suppose a participant is shown 50 words, what is the probability that the participant will correctly recognize 45 out of 50 words? You are given that the probability of correctly recognizing a word is 0.9.

You are given a probability assigner function $f(k, n, p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$, where k is number of correctly recognized words, n is total number of words, and p is underlying probability of correctly recognizing a word.

2.2 Suppose in a small city, 10 road accidents happen on average in a single day. The probability of k number of road accidents in a day is given by the probability mass function $f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$ where $\lambda = 10$.

(a) What is the probability that zero road accidents in a day?

(b) What is the probability of ocurrence of more than 7 but less than 10 road accidents in a day?

(c) Draw the graph of the probability mass function $f(x)$.
(Hint: Choose a range of possible values for X say 0,1,2,...,50; compute the probabilities for each value of X using the formula for $P(X = x)$; and then, plot the graph with probabilities on y-axis and X on x-axis.)

Part 3: Continuous random variables

3.1 Suppose a random variable X is normally distributed. The probability density function of the normal distribution is given by $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

(a) Find the probability density of obtaining $x = 0$, given that $\mu = 1, \sigma = 1$.

(b) Find the probability density of obtaining $x = 1$, given that $\mu = 0, \sigma = 1$.

(c) You are given
The probability that the outcome occurs between x_1 and x_2: $P(x_1 \leq X \leq x_2) = 0.3$
The probability that the outcome occurs between x_1 and x_3: $P(x_1 \leq X \leq x_3) = 0.45$
Find the probability that the outcome occurs between x_2 and x_3.

Part 4: The likelihood function

Suppose a random variable X has the probability density function $f(x, \theta)$ where θ is a parameter of the probability density function and x is a value of the random variable X. You can write:

$X \sim f(x, \theta)$

The probability density function (PDF) $f(x, \theta)$ tells you the probability density of generating an outcome x when the value of θ is known/assumed. For example, if you know/assume $\theta = 2$, you can calculate the probability density for different values of the random variable such as $X = 5$, $X = 3$, $X = 100$. Basically, the PDF is viewed as a function of x when θ is fixed.

However, you can also view the PDF in a different manner. You can calculate the probability density of obtaining a given, fixed outcome x for different values of θ. That is, the PDF can be viewed as a function of θ when x is fixed. This alternative characterization of the PDF is called **the likelihood function**.

The likelihood function is a function that maps the values of the parameter θ to probability densities, when the sample x is taken as a fixed, observed quantity.

In sum, the PDF is a function of x where θ is assumed to be fixed; the likelihood function is a function of parameter θ when the sample x is fixed/known.

The likelihood function is often represented by $\mathcal{L}(\theta|x)$:

$\mathcal{L}(\theta|x) = f(x, \theta)$ when x is fixed

Use the above information to do the following exercise.

4.1 In a visual word recognition experiment, a participant has to recognize whether a string shown on screen is a meaningful word (e.g., "book") or a non-word (e.g., "bktr"). The participant is asked to answer "yes" if the shown string is a meaningful word, and "no" if it is a meaningless non-word. Suppose a participant is shown 5 strings on the screen one by one. The time taken by the participants to recognize each string is shown below (in milliseconds):

Recognition time for 5 strings: $303, 443, 220, 560, 880$

Suppose the random variable X represents the string recognition times.

A researcher proposes a hypothesis that the string recognition times are generated by the probability density functions $f(x, \mu)$:

$X \sim f(x, \mu)$

such that,

$f(x, \mu) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2}}$

(a) Plot the graph of the likelihood function with respect to values of μ, assuming that x is fixed to 220.

(Hint: Choose a range of values for μ and plug those values in the function $f(x, \mu)$ along with $x = 220$ to get probability densities; plot a graph with μ on x-axis and the corresponding values of $f(x, \mu)$ on y-axis.)

(b) Graph the likelihood function when x is (fixed as) the observed sample of recognition times i.e., $303.25, 443, 220, 560, 880$.

(Hint: The probability density for any $\mu = \mu_1$ and $x = [x_1, x_2, x_3, ..., x_n]$ will be $f(x_{1:n}, \mu_1) = \frac{1}{\left(\prod_{i=1}^n x_i\right)\left(\sqrt{2\pi}\right)^n} e^{-\frac{\sum_{i=1}^n (\log x_i - \mu_1)^2}{2}}$.)

(c) For what value of μ, the likelihood (probability density) of obtaining the observed sample $303, 443, 220, 560, 880$ is maximum? You do not need to be precise, you can tell the approximate value.