Lecture 19 Worksheet

Chrysafis Vogiatzis

Every worksheet will work as follows.

1. You will be entered into a Zoom breakout session with other students in the class.
2. Read through the worksheet, discussing any questions with the other participants in your breakout session.
   - You can call me using the "Ask for help" button.
   - Keep in mind that I will be going through all rooms during the session so it might take me a while to get to you.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students in a breakout session is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see Compass) to submit your work.

Worksheet 1: A simple Bayesian estimation

As a reminder, to get the Bayesian estimators, we:

1. identify our prior probabilities/distribution: this quantifies what we think the unknown parameters are *before* observing any sample.
2. build the likelihood function $L(\theta)$ by multiplying the probability mass function (for discrete) or the probability density function (for continuous) for each of the sample values (**exactly what we did last time for MLE**).
3. compute the posterior probabilities/distribution by multiplying the priors with the likelihood.
4. find the maximum posterior probability or find the point where the posterior distribution attains its maximum.
   - This latter part can be done by either comparing all probabilities and picking the biggest one or
   - by getting the derivative(s) of the posterior distribution for each of the unknown parameter(s) and setting it to 0

Problem 1: Creating a table

For smaller problems, it may be useful to create a table. In the table, we store the prior probabilities (based on past information or prior beliefs), the likelihood functions (based on the sample), and computer the posterior probabilities. Let us see this in an example.

A school can be in one of three categories: $R1, R2, R3$. $R1$ schools place 90% of their graduates in good jobs, $R2$ schools place 75% of their graduates in good jobs, and $R3$ schools place 50% of their graduates in good jobs. You may assume that each school has a $\frac{1}{3}$ chance of appearing.

You have been observing a school and have found that 82 of their last 100 graduates have been placed in good jobs. Is the school an $R1$, $R2$, or $R3$ school? Equivalently, if p is the probability of placing a graduate in a good position, what is $\hat{p}$ for the specific school begin observed?

Answer to Problem 1.

	priors $P(p)$	likelihood $L(X\|p)$	posterior $P(p) \cdot L(X\|p)$
$p = 0.9$			
$p = 0.75$			
$p = 0.5$			

Problem 2: Normalizing the result

Based on your calculations, you probably noticed that the posterior probabilities are very small. Can you normalize them so as to answer the question: "how certain are you that the school is an R1/R2/R3 institution?" [1]

[1] Normalizing entails getting each end result and dividing by the summation of all of your results.

Answer to Problem 2.

Worksheet 2: Streaming services and their data (Bayesian remix)

We turn one last time to the data from the TV streaming giant. As a reminder, they have given you data on the time people spend during a session (continuous); as well as the number of episodes people watch during a session (discrete number). In this worksheet, we focus solely on the number of episodes streamers watch per session.

	Session #									
	1	2	3	4	5	6	7	8	9	10
Time (in hours)	0.75	1.20	1.33	0.97	0.80	1.43	0.87	1.41	1.09	1.05
# of episodes	1	3	3	3	2	5	1	5	3	1

Problem 3: Poisson with prior beliefs

The streaming giant believes that there are three types of customers. For all three types the number of episodes they watch is Poisson distributed; they do have different rates λ, though. The three types of customers are:

- 75% of their clientele watches a number of episodes that is Poisson distributed with $\lambda = 1$.
- 10% of their clientele watches a number of episodes that is Poisson distributed with $\lambda = 3$.
- 15% of their clientele watches a number of episodes that is Poisson distributed with $\lambda = 4$.

Using this new information, and assuming that all data has been collected from one type of customers alone, what is the Bayesian estimator for λ?

Answer to Problem 3.

Problem 4: Poisson with more general prior beliefs

Here we go in the general, continuous case. What if the streaming giant was wrong and customers are not discrete (that is $\lambda = 1, 3$, or 4), but are instead continuous (λ is *anything* in the 1 to 4 range)? For this problem, you may assume that λ is uniformly distributed between 1 to 4. What is the Bayesian estimator for λ in this case? [2]

[2] You will need again the fact that the pdf of the uniform distribution is $\frac{1}{b-a}$, where a and b are the lower and upper bounds of the uniform.

Answer to Problem 4.

We may summarize this result as follows. When presented with discrete cases, we will calculate posterior *probabilities* (normalize them, if we prefer) and report the maximum among them. In the continuous case, we will calculate the posterior *distribution* and find the maximizer (potentially by setting the derivative equal to 0).

Worksheet 3: Coins

Problem 5: Step-by-step

Assume you carry with you three coins with probability of Heads $p = 0.25, 0.5, 0.75$. You pick a coin and you flip it. Which coin do you (believe you) have picked to flip if: [3]

[3] Assume the events presented later come in sequence.

Answer to Problem 5.

the first coin comes up Heads?

the second coin comes up Heads?

the third coin comes up Tails?

the fourth coin comes up Tails?

the fifth coin comes up Tails?

Now, let us run into one of the interesting aspects of Bayesian estimation...

Problem 6: Coins again

You still have an unfair coin at your disposal that brings Heads with unknown probability p. This time though, your coin is not one of three like earlier. Instead, the probability p is uniformly distributed between 0.4 and 0.6.

You decide to run an experiment to estimate p. You will toss the coin as many times as necessary to get Heads and record the number of times until Heads are observed. [4] You have obtained $N_1 = 5, N_2 = 4, N_3 = 5$.

[4] This is... geometric, no?

With this in mind, what is the Bayesian estimator for p?

Answer to Problem 6.

This is interesting! If our maximizer is above the largest allowable value, or below the smallest allowable value, then we need to pick the largest/smallest one! You may also show this graphically, if you prefer, to see what happens.

Worksheet 4: Continuous case only

What if both the distribution we are estimating is continuous **and** the distribution of the parameter is continuous at the same time?

Problem 7: Normal and exponential

The time between two consecutive vehicles passing through an intersection is exponentially distributed with unknown rate λ; however, we do know that λ is normally distributed with $\mu = 2$ per minute and variance $\sigma^2 = 0.25$. In essence, this states that typically 2 vehicles would pass every minute, but this can go lower to 1 every 2 minutes (when it is less busy) and higher to 7 every 2 minutes (when it is much busier).

A new day has just begun and we would like to see how many vehicles to expect in the next 2 hours. During the first 5 minutes, we saw the first vehicle appear in $T_1 = 0.5$ minutes, $T_2 = 1$ minutes, $T_3 = 2$ minutes, $T_4 = 1.5$ minutes. Recall that these are times between two consecutive vehicles.

With this in mind, what is the Bayesian estimator for λ?

Answer to Problem 7.