# Homework 1

Fabrizio Gandossi

October 17, 2024

**Exercise 1.** Consider the coin tossing example, discussed in the first lecture. Simulate 1000 tosses of the coins, setting H = 0.3. Consider a uniform prior and update the posterior at each toss. Plot the resulting posterior after 1, 50, 100, 300, 700, 1000 tosses. Repeat the simulated experiment by setting a Gaussian prior centered in H = 0.5, with standard deviation $\sigma$= 0.1. Do both posteriors converge a similar distribution in the end? What does that mean? Which posterior converges faster and why?

*Solution.* To solve the exercise I make use of the Bayes' Theorem, but first I need to generate a random sequence of 1000 tosses, and then I write how many tosses I'm interested in. I just take the first tosses of the array and not chosen by random because it's statistically the same thing.

```
#importing necessary packages
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

#parameters
H = 0.3  #probability of heads
N = 1000  #number of tosses
obs_tosses = [0, 1, 50, 100, 300, 700, 1000] #number of tosses for
   plotting
```

Then I simulate the tosses following a binomial distribution and I define the range of the possible MAP.

```
#simulating tosses (1 = head, 0 = tail)
tosses = np.random.binomial(1, H, N)

#possible values of H
h_range = np.linspace(0, 1, 1000)
```

I build the posterior by multiplying the likelihood $L = (h_{\text{values}})^{(\text{Number of heads})} * (1 - h_{\text{values}})^{(N-\text{Number of heads})}$ with a uniform prior, which in this case I can omit and then plot the distribution with the requested number of tosses and normalizing it with the peak at 1.

```
#plotting all the distriubtions
for i in range(len(obs_tosses)):
```

```
    N_heads = np.sum(tosses[:obs_tosses[i]])
    posterior = (h_range ** N_heads) * ((1 - h_range) ** (obs_tosses[i] - N_heads))
    plt.plot(h_range, posterior/np.max(posterior), label=f"number of tosses: {obs_tosses[i]}", linewidth = 1)

#plotting the real value dashed
plt.axvline(x=0.3, color='grey', linestyle='--', label='true H=0.3', alpha = 1)

plt.legend()
plt.xlabel("H")
plt.ylabel("P(H|data)")
plt.title("Distriution with a uniform prior")
plt.show()
```

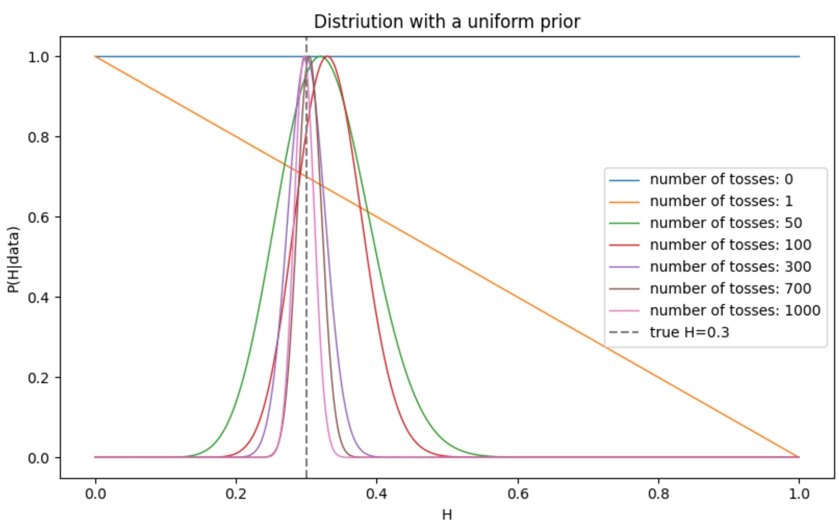


Figure 1: Different distributions with a uniform prior.

Computing the same thing but using a gaussian distribution as a prior centered on $\mu = 0.5$ and with $\sigma$=0.1 we obtain the following posteriors

```
    ### now i just do the same but putting a gaussian prior

for i in range(len(obs_tosses)):
    N_heads = np.sum(tosses[:obs_tosses[i]])
    posterior = (h_range ** N_heads) * ((1 - h_range) ** (obs_tosses[i] - N_heads)) * norm.pdf(h_range, loc=0.5, scale=0.1)
    plt.plot(h_range, posterior/np.max(posterior), label=f"number of tosses: {obs_tosses[i]}", linewidth = 1)
```

```
plt.axvline(x=0.3, color='grey', linestyle='--', label='true H=0.3',
    alpha = 1)
plt.legend()
plt.xlabel("H")
plt.ylabel("P(H|data)")
plt.title("Distriution with a gaussian prior")
plt.show()
```

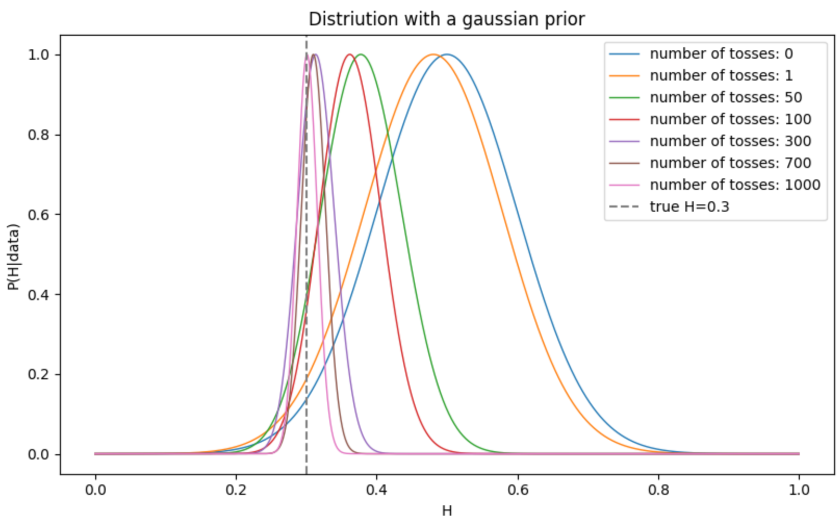


Figure 2: Different distributions with a gaussian prior.

I can now compute which of the two distribution converge fastest by plotting how the value of the maximum a posteriori estimate changes with the number of coin tosses and we see that the uniform prior approaches the real value faster than the gaussian one.

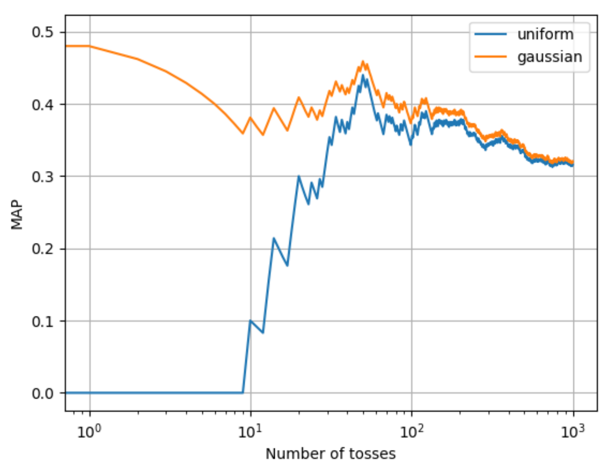


Figure 3: Trend of the MAP.

□

**Exercise 2.** Politician A makes a statement about some issue you knew nothing about before. Let's call such proposition S and assume your starting prior on S is uniform with 0.5 probability of S being either true or false. Update your probability of S being true, knowing that you trust Mr. A to tell the truth with probability prob($A_T$) = 4/5. At this point Mr B - another politician - declares that he agrees with Mr A on S being true. You trust Mr. B much less, and believe that the probability of him to lie is prob($B_T$) = 3/4. What is your final degree of belief in proposition S?

*Solution.* For the first question we use again the Bayes' theorem starting with the following definition:

$$P(S|A_T) = \frac{P(A_T|S) \cdot P(S)}{P(A_T)} \tag{1}$$

where:

- $P(S|A_T)$ is the posterior probability, or the probability of S to be true given the statement of A.
- $P(A_T|S)$ is the likelihood, or the probability of A saying that S is true given that is really true.
- $P(S)$ is the prior probability that S is true.
- $P(A_T)$ is the normalization factor, and it's the probability of A saying that S is true regardless of being true or false.

In our starting point we chose $P(A_T|S) = 4/5$, $P(S) = 0.5$ and $P(A_T) = P(A_T|S) \cdot P(S) + P(A_T|\bar{S}) \cdot P(\bar{S}) = 0.5$.

So by putting everything together we get:

$$P(S|A_T) = \frac{0.8 \cdot 0.5}{0.5} = 0.8 \tag{2}$$

and this is the probability of S being true given that A told that is true.
To continue with the second point we can make the same reasoning but using as a prior the posterior that we've just found, so that the equation becomes:

$$P(S|A_T, B_T) = \frac{P(B_T|S) \cdot P(S|A_T)}{P(B_T)} \tag{3}$$

We just have to compute the normalization factor

$$P(B_T) = P(B_T|S) \cdot P(S|A_T) + P(B_T|\bar{S}) \cdot P(\bar{S}|A_T) = (0.25 \cdot 0.8) + (0.75 \cdot 0.2) = 0.35 \tag{4}$$

Putting everything in the equation above we obtain:

$$P(S|A_T, B_T) = \frac{0.25 \cdot 0.8}{0.35} \approx 0.57 \tag{5}$$

And this is our final belief that the statement S is true given that both politician A and B told that S is true.

□

**Exercise 3.** You are tested for a dangerous disease named "Bacillum Bayesianum" (BB). You test positive to BB. You know that the general incidence of BB in the population is 1%. Moreover, you know that your test has a false negative probability of 5% (false negative: you have BB but the test scores negative), and a false positive rate also of 5% (false positive: you do not have BB, but the test scores positive). What is the probability that you have actually contracted BB?

*Solution.* This is plain and simple application of the Bayes' Theorem, which we write in the form:

$$P(BB|test_p) = \frac{P(test_p|BB) \cdot P(BB)}{P(test_p)} \tag{6}$$

where:

- $P(BB|test_p)$ is the posterior probability, or the probability of having BB after testing positive.
- $P(test_p|BB)$ is the likelihood, or the probability of tasting positive given that you have BB.
- $P(BB)$ is the prior probability to have BB, or the general incidence.
- $P(test_p)$ is the normalization factor, or the probability to have a positive test regardless having or not BB.

The only thing to compute is $P(test_p)$ and, as before

$$P(test_p) = P(test_p|BB) \cdot P(BB) + P(test_p|\overline{BB}) \cdot P(\overline{BB}) = (0.95 \cdot 0.01) + (0.05 \cdot 0.99) = 0.059 \tag{7}$$

and putting everything together we obtain

$$P(BB|test_p) = \frac{0.95 \cdot 0.01}{0.059} \approx 0.161 \tag{8}$$

and this is the probability that we've actually contracted BB. □