Facultat de Física de la Universitat de Barcelona
Probability and Statistics

Task 4

Javier Castillo Uviña
jcastiuv7@alumnes.ub.edu

1.

From theory we know Bayesian inference works pretty well with a small dataset, using the following expression we are about to find the posterior expression. Maximizing we will find the best estimator for the parameter θ.

$$p(\theta \mid y) = \frac{p(y \mid \theta) \cdot p(\theta)}{p(y)}$$

- $p(\theta)$ is the prior probability. We used three different priors:
  1. $p(\theta) = 1$ following a uniform distribution between $[0, 1]$.
  2. $p(\theta) = 2\theta$ following a linear distribution between $[0, 1]$.
  3. $p(\theta) = -2\theta + 2$ following a inverse linear distribution between $[0, 1]$.
- $p(y \mid \theta)$ is the probability density of observing y given our hypothesis θ, and is called the likelihood. In our case this y means we observed n_h heads and n_t toes, so we obtained $p(n_h, n_t \mid \theta) = \theta^{n_h}(1-\theta)^{n_t}$.
- $p(y)$ is the normalization and refers to the total probability cases we are working with. In this case our normalization term is $p(n_h, n_t) = \int_0^1 p(\theta)p(n_h, n_t \mid \theta)d\theta$ using the different priors for each case.

In the first graphic 1 we observe the posteriors for the different priors for $n_h = 3 n_t = 7$, each one has its own best estimator for θ. Uniform takes $\theta = 0.3$ value; linear takes $\theta = 0.36$ and the final one inverse linear $\theta = 0.27$. We can deduce θ is the head probability, the linear prior favor big θ that's why it is the biggest best estimator value. The opposite casuistry occurs for the linear.
The second graphic 2 represents the same analysis for $n_h = 8 n_t = 2$ and the best estimator values are: uniform takes $\theta = 0.8$; linear $\theta = 0.82$ and inverse linear $\theta = 0.73$.

$\boldsymbol{MLE\ approach}$
We construct the likelihood function, which is already given by the sentence:

$$L(\theta) = \theta^{n_h}(1-\theta)^{n_t}$$

And know we have to maximize $L(\theta)$ with respect to θ and we will obtain.

$$\hat{\theta} = \frac{n_h}{n_h + n_t}$$

For the two different cases we obtein just a value each:

- $n_h = 3 \; n_t = 7 \rightarrow \theta = 0.3$
- $n_h = 8 \; n_t = 3 \rightarrow \theta = 0.8$

Observations:

- MLE has the same behaviour that the Bayesian inference for a uniform prior.
- Non Bayes either MLE methods gave a value even similar $\theta = 0.5$, so the coin is biased for each case.

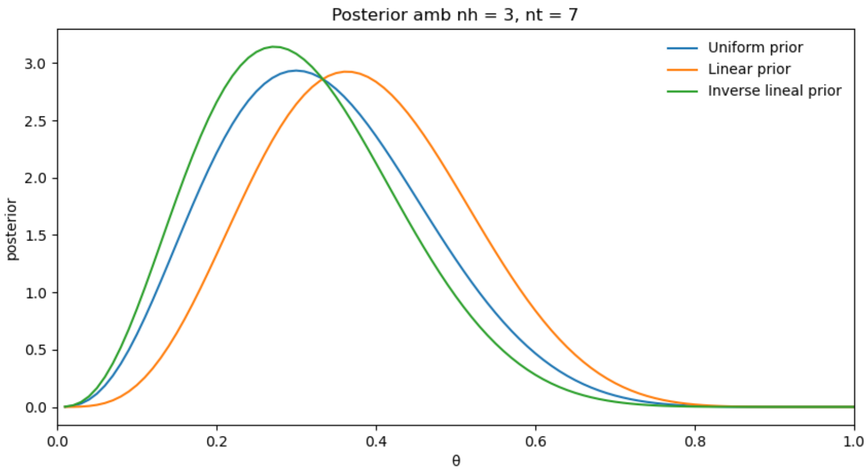

Figure 1 – Posterior object for the $n_h = 3 \; n_t = 7$ case for different priors.

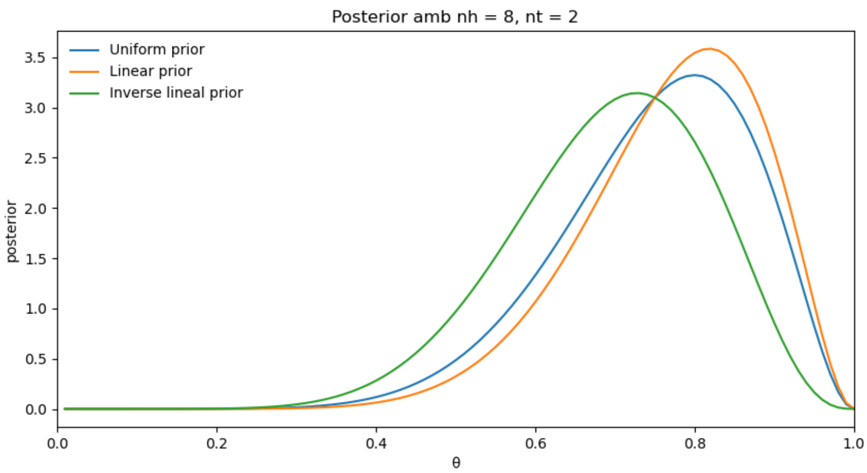

Figure 2 – Posterior object for the $n_h = 8 \; n_t = 3$ case for different priors.

2.

In order to confection a correct test there are 4 steps to follow:

1. Sate a hypothesis H_0 called the $null\ hypothesis$. Now we want to prove if this null hypothesis is solid with the data, so we construct H_1 which contradict the first one. They have to fulfill:

- $H_0 \cap H_1 = \emptyset$
- $H_0 \cup H_1 = \Omega$

So we are choosing H_0: the mean data set value $\overline{x} = 25$. On the other hand H_1: the mean data set value $\overline{x} \neq 25$.

2. Select the criterion to accept this new hypothesis over the null hypothesis. We select the significance $\alpha = 0.05$ as the sentence says.
3. Use a statistical test such Z-test or T-test. We will use both because the results are different.
4. We establish a $p - value = 1 - P(Z \leq z^{test})$. Case $p < \alpha$ we reject the hypothesis. Case $p \geq \alpha$, the hypothesis is failing to reject.

We must observe that this is a bidirectional test, so our significance $\alpha = 0.025$.

$\boldsymbol{T - Test}$

$$z_T = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{26.68 - 25}{2.09/\sqrt{10}} = 2.542 \Rightarrow p = 1 - 0.9945 = 0.0055 < \alpha = 0.025$$

$\bar{x}$ is the experimental data set mean, μ is the theoretical mean, s is the set standard deviation and n is the set length. In this case we have to reject the null hypothesis due to our algorithm.

$\boldsymbol{Z - Test}$

$$z_T = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{26.68 - 25}{4/\sqrt{10}} = 1.328 \Rightarrow p = 1 - 0.9082 = 0.0918 > \alpha = 0.025$$

σ is the theoretical standard deviation. In this other case we fail to reject the null hypothesis.
Due to our set size $n = 10$ versus a what is considered a big sample sized set of $n = 30$ it is preferable to use the T-Test. So we can reject the null hypothesis with a significance of 5%.