Exercise sheet 6

Exercise 6 - 1

Consider the following coin toss experiment:

- A large number n of coin tosses are performed and the results are stored in a data vector $d^{(n)} = (d_1, \ldots, d_n) \in \{0,1\}^n$, where 0 and 1 represent the possible outcomes head and tail.
- Individual tosses are independent from each other.
- All tosses are done with the same coin with an unknown bias $f \in [0,1]$; i.e., $\mathcal{P}(d_i|f) = f^{d_i}(1-f)^{1-d_i}$.

Assume that a fraction $\bar{f}$ out of the n coin tosses yielded head.

a) Derive the Gaussian approximation of the PDF $\mathcal{P}(f|d^{(n)})$ around its maximum. – You can use a saddle point approximation; i.e., identify the maximum, and taylor-expand the (negative) logarithm of $\mathcal{P}(f|d^{(n)})$ around it up to second order in order to identify the variance of the Gaussian (3 points).

b) Use this Gaussian approximation to derive an approximation for $\mathcal{P}(d_{n+1}|d^{(n)})$.
Hint: You can assume that the Gaussian distribution is narrow enough such that the integration boundaries $[0,1]$ can be replaced by $(-\infty, \infty)$ (2 points).

c) Now calculate the exact posterior mean for $\bar{f}$ and the exact expression for $\mathcal{P}(d_{n+1}|d^{(n)})$ (2 points).

Note: $\int_0^1 \mathrm{d}x\, x^\alpha(1-x)^\beta = \Gamma(\alpha+1)\Gamma(\beta+1)/\Gamma(\alpha+\beta+2)$, where Γ is the Gamma function.

Exercise 6 - 2

Let $P(s) = \mathcal{G}(s, S)$ with $s = (s_1, ..., s_n)^t$ be a real multivariate zero-centered Gaussian with Covariance $\langle ss^t \rangle = S$. We would like to fit another Gaussian distribution $P'(s) = \mathcal{G}(s, S')$ to it that has a diagonal covariance matrix $S'_{ij} = \delta_{ij}\sigma_i$. Here δ_{ij} denotes the Kronecker delta.

a) What is the optimal approximating Gaussian $P'(s)$ to $P(s)$, as parameterized by σ, obtained through minimizing the loss

$$\sigma = \arg\min_\sigma \mathrm{KL}(P(s), P'(s)) \ ?$$

(2 points)

b) What is the least updating fit of $P'(s)$ to $P(s)$, as obtained through

$$\sigma = \arg\min_\sigma \mathrm{KL}(P'(s), P(s)) \ ?$$

(2 points)

c) Let

$$S = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Visualize the Gaussian $\mathcal{G}(s, S)$ as well as the two different ways to fit a diagonal Gaussian to it introduced above. To visualize them, use a computer plotting samples of each of the three distributions. You can draw a sample from a multivariate Gaussian distribution by applying the square root of its covariance matrix to a white noise sample. (optional)

Exercise 6 - 3

Assume the measurement of a signal s, which yields the data d, leads to

$$\begin{aligned}\mathcal{P}(s) &= \mathcal{G}(s, S)\\ \mathcal{P}(s|d) &= \mathcal{G}(s-m, D).\end{aligned}$$

a) Calculate the amount of information in terms of entropy one gains via the measurement (3 points).

b) Now assume that your signal prior in the above measurement was flat. How much information in terms of entropy does one gain via the measurement? Explain the result (1 point).

This exercise sheet will be discussed during the exercises.
Group 01, Wednesday 18:00 - 20:00, Theresienstr. 37, A 449,
Group 02, Thursday, 10:00 - 12:00, Theresienstr. 37, A 249,
Group 03, Thursday, 16:00 - 18:00, Theresienstr. 37, A 449

https://wwwmpa.mpa-garching.mpg.de/ ensslin/lectures/lectures.html