Assignment 8

1. (**10**) Assume you are flipping a biased coin and the outcome of 100 flips is 75 Heads and 25 Tails. Compute the Maximal Likelihood Estimation for the Head probability p_0 and its error δp (the error is defined such that the probability of p to lie in the interval $[p_0 - \delta p, p_0 + \delta p]$ is 95%).

2. (**10**) Let x_1 and x_2 be normal random variables, $x_1 \sim (0, \sigma_1)$ and $x_2 \sim (0, \sigma_2)$.
   - Find the distribution function of $y = x_1 + x_2$.
   - Compute the integral $Q(y) = \int dx P(x|\sigma_1) P(y - x|\sigma_2) dx$, where $P(x|\sigma)$ is probability density of $x \sim P(0, \sigma)$. What is probabilistic interpretation of tis integral?

3. (**20**) Multivariate normal distribution over vectors of length k is defined as follows:

$$\mathcal{N}(\mathbf{x}|\mathbf{\Sigma}) = (2\pi)^{-k/2} \det(\mathbf{\Sigma})^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{\Sigma}^{-1} \mathbf{x}\right). \tag{1}$$

   Your goal is to compute the following integral:

$$I(\mathbf{x}, \mathbf{\Sigma_1}, \mathbf{\Sigma_2}) = \int d^k \mathbf{y} \mathcal{N}(\mathbf{x} - \mathbf{y}|\mathbf{\Sigma}_1) \mathcal{N}(\mathbf{y}|\mathbf{\Sigma}_2) \tag{2}$$

   in two ways. Before proceeding to the multivariate case, you may consider $k = 1$ for simplicity.
   - (**5**) Having this in mind the probabilistic interpretation of this integral, the result of this integration can be written down immediately (consult Eq. (1)).
   - (**15**) The integral $I(\mathbf{x}, \mathbf{\Sigma_1}, \mathbf{\Sigma_2})$ can be reduced to the standard integral:

$$\int e^{-\frac{1}{2}x^\mathsf{T}\mathbf{A}x + B^\mathsf{T}x} d^k x = \sqrt{\frac{(2\pi)^k}{\det A}} e^{\frac{1}{2}B^\mathsf{T}\mathbf{A}^{-1}B}. \tag{3}$$

     Use this fact to re-derive the result of the previous item.

4. (**15**) Consider the model, described by the following state and observation equation

$$x_t = x_{t-1} + u_{t-1} + \epsilon_{t-1}, \quad z_t = x_t + \eta_t \tag{4}$$

   with $P(\epsilon_t) = \mathcal{N}(0, \sigma)$, $P(\eta_t) = \mathcal{N}(0, \rho)$. As was shown on the lecture, the max-likelihood estimate of the hidden state conditioned on all information observed by the moment t reads:

$$\hat{x}_t = K_t z_t + (1 - K_t)(\hat{x}_{t-1} + u_{t-1}). \tag{5}$$

   Compute $\lim_{t\to\infty} K_t$ and $\lim_{t\to\infty} \left\langle (\hat{x}_t - x_t)^2 \right\rangle$.

5. (**40**) Consider flipping two biased (a priori, the bais of each coin is uniformly distributed on $[0, 1]$) coins. We repeat the following experiment 3 times: (i) Pick one of the two coins at random (ii) Flip the selected coin 20 times. The results of the three experiments are as follows: `TTTTHHHTTTHTTTHHTTTT`, `HHHHHHHHHHHHHHTTTHHHH`, `HHHTHHHHHHTHTTHTHHHH`.
   - (**10**) Compute the maximum likelihood estimate of the coins biases p_1 and p_2.
   - (**10**) Estimate a posteriori probability that biases of the coins are approximately the same: $P(|p_1 - p_2| < 0.1)$.
   - (**20***) Estimate a posteriori probability that in each of the three experiments above exactly the same coin was selected for flipping.

6. (**35**) Consider the model, described on the lecture:

$$x_t = x_{t-1} + u_{t-1} + \epsilon_{t-1}, \quad z_t = x_t + \eta_t \tag{6}$$

   where z_t is observed state, x_t is hidden state and, for simplicity, $u_t \equiv u$. Assume that the measurement and the control noise have the following distributions:

$$P(\epsilon_t) = \mathcal{N}(0, \sigma), \quad P(\eta_t) = \mathcal{N}(0, \rho = 2\sigma). \tag{7}$$

   Compute, analytically, the following distributions:

(a) (**5**) $P(x_0|z_0)$: distribution of x_0 after the first 1 observation

(b) (**10**) $P(x_0|z_0, z_1)$: distribution of x_0 after the first 2 observations

(c) (**20***) $P(x_1|z_0, z_2)$: distribution of x_1 after the first 3 observations, when the second observation is missing.

Note that each of these distributions is Gaussian, so its enough to specify mean and variance in each case. Feel free to use Eq. (3) (`Mathematica` or `sympy` may also be useful for algebraic manipulations.)