TSKS15 Additional Problems

Erik G. Larsson

September 12, 2017

Problems 1-3 are intended as repetition of the probability class. Students who master Bayesian probability theory do not need to work on these problems.

1. Fanny tells you he has two sisters, Ann and and Britney.

   a) What is the probability that Fanny is older than Britney?

   b) Fanny tells you that she is older than Ann. What is the probability that Fanny is older than Britney?

2. The disease X has prevalence 1% in the population and Joe is a randomly selected individual in that population.

   Joe tests for the disease. The test is 95% reliable in the sense that 95% of the people with the disease get a positive test result, and 95% of those free of the disease get a negative result.

   Suppose Joe tests positively. What is the probability that he has the disease?

3. (Three doors problem.) You participate in game where there are three doors are labelled A, B and C. Behind one of them is a prize (you don't know which one), but behind the other two there is nothing. You get to select one door. The host then opens *one of the other two doors*, but one that does not have a prize behind it – such that the location of the prize is not revealed.

   At this point, you may either stick to your original choice of door, or switch to the other (unopened) door. What should you do? Stick to your original choice, switch – or does it not matter?

4. (After David MacKay.) You are given a coin, and the task is to determine whether the coin is fair (H_0) or not (H_1).

   a) In an experiment you decide to flip the coin a fixed number (n) of times and record how many heads (h) and how many tails (t) you got ($n = h + t$). You then compute the probability p_e that *if the coin were fair, you would obtain an outcome as extreme, or more extreme, than this*. If p_e is less than a predefined, fixed "plausibility" threshold p_0 (take 5% in this problem) then you determine the coin is biased (H_1), otherwise you say it is fair (H_0).

   Suppose you decide to take $n = 12$ (flip the coin 12 times) and you obtain the sequence `hhhthhhhthht`. Do you decide on H_0 or H_1?

   b) In a different experiment you decide instead to flip the coin until a fixed number, say t, of tails occur. You then record how many times (n) you had to flip the coin for this to happen. As before you then compute the probability p_e that if the coin were fair, you would obtain an outcome that is at least as extreme as this, and decide on H_1 if p_e is less than the level 5%.

   Suppose you decide to take $t = 3$ and obtained the string above,

   `hhhthhhhthht`,

   i.e. the 3rd tail occurred after $n = 12$ tosses. Do you decide on H_0 or H_1?

   c) Do the experiments (a) and (b) lead to the same conclusion? If not, why? Which experiment (a) or (b) do you think is most "fair"? Which one do you prefer and why?

   d) Design a Bayesian test and compute the a posteriori likelihood ratio explicitly. You may make any assumptions you wish, e.g., as in the lecture.

   Hints: In a) you need to consider the distribution of t, for n fixed. In b), you need to consider the distribution of n, for t fixed.

5. A single sample is observed. The model is

$$H_0 : y = -1 + w$$

$$H_1 : y = 1 + w$$

   where $w \sim N(0, \sigma^2)$. Bayesian theory is used. It is known that $P(H_0) = 1/4$ and $P(H_1) = 3/4$.

   a) Give the optimal min-probability(error) Bayesian detector.

b) For the optimal detector, what happens to the threshold when $\sigma \to 0$? What happens to the error probability when $\sigma \to 0$? Explain intuitively.

c) For the optimal detector, what happens to the threshold when $\sigma \to \infty$? What happens to the error probability when $\sigma \to 0$? Explain intuitively.

d) Write out the suboptimal detector that would result if the prior were ignored and one just set $P(H_0) = P(H_1) = 1/2$ in the decision rule.

e) For the suboptimal detector, does the threshold depend on σ? Why, or why now?

f) For the suboptimal detector, what happens to the error probability when $\sigma \to 0$? Explain intuitively.

g) For the suboptimal detector, what happens to the error probability when $\sigma \to \infty$? Explain intuitively.

h) Sketch, qualitatively, the error probability as function of σ for the optimal and suboptimal detector.

6. For some problems, the Cramér-Rao bound does not exist because the Fisher information matrix $\boldsymbol{I}$ is singular (not invertible). Such problems are said to be **unidentifiable**.

   a) Consider the estimation of A, B and C in the following model

   $$y_n = A\cos(2\pi fn) + B\sin(2\pi fn) + C\cos(2\pi fn + \pi/4) + w_n, \quad n = 0, \ldots, N-1$$

   where w_n are independent $N(0,1)$ and f is a known frequency that is an integer multiple of $1/N$. Compute $\boldsymbol{I}$. Is this problem unidentifiable or not? If it is, intuitively what is the reason?

   b) Is any of the examples in Chapter 3 of Kay's book (part I) unidentifiable? Any examples in my lecture?

   c) If you encounter an unidentifiable problem in an engineering application, what would you do?

7. A linear-algebra problem that repeatedly occurs in this course is the following:

   $$\boxed{\text{minimize} \quad \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta}\|}$$

   Its solution is, assuming $\boldsymbol{H}$ has linearly independent columns,

   $$\boxed{\boldsymbol{\theta} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{y}}$$

a) Explain each step in the following sequence of (in)equalities:

$$
\begin{aligned}
\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta}\|^2 &= \|\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta}) + (\boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T)(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta})\|^2 \\
&= \|\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta})\|^2 + \|(\boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T)(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta})\|^2 \\
&= \|\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta})\|^2 + \|(\boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T)\boldsymbol{y}\|^2 \\
&\geq \|\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\theta})\|^2 \\
&= \|\boldsymbol{H}((\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{y} - \boldsymbol{\theta})\|^2
\end{aligned}
$$

b) Explain why the chain of (in)equalities in (a) implies the solution given above to the minimization problem, $\boldsymbol{\theta} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{y}$.

c) What other ways of solving this minimization problem do you know of?

d) What happens if the columns of $\boldsymbol{H}$ are not linearly independent?