Example exam

DIRECTIONS: ALLOW 3 HOURS. WRITE YOUR NAME AND STUDENT NUMBER AT THE TOP OF EVERY PAGE OF YOUR SOLUTIONS. Please explain clearly all of the steps that you used to derive a result. Please make certain that your handwriting is readable to someone besides yourself.

1. Answer the following open questions: **7 points/question**

   (a) Using Cox's product and/or sum rules, derive the mathematical formula that describes Bayes' theorem and give the name of each term in the formula.

   (b) Describe and compare parameter estimation and model comparison. What role does the evidence play in each process?

   (c) In the context of parameter estimation, describe the difference between a scale parameter and a location parameter.

   (d) Given a continuous uniform distribution function $\text{prob}(x) = \frac{1}{3}$ for $1 \le x \le 4$ and $\text{prob}(x) = 0$ otherwise:

      i. Calculate the expectation value $\langle 3\sqrt{x} + 6 \rangle$

      ii. Calculate the expectation value $\langle x \rangle$ and the variance of x, $\text{Var}(x)$.

   (e) Explain what the Kolmogorov–Smirnov test is used for and how it works.

   (f) Given a one-dimensional probability density function, describe how you would find the 95% confidence interval.

   (g) Given the joint posterior distribution of X and Y, $\text{prob}\,(X, Y \mid \{\text{data}\}, I)$, write down the definition of the covariance of X and Y. What does it mean when the covariance is equal to zero?

   (h) Describe what a singular value decomposition is and include the formula for the decomposition. How can a singular value decomposition be used in multi-variate parameter estimation, and what is its advantage over other methods?

   (i) For solving a two-parameter problem with parameters X and Y, the following equation can be defined:

   $$Q(X,Y) = (X - X_0 \;\; Y - Y_0) \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} X - X_0 \\ Y - Y_0 \end{pmatrix}$$

   with

   $$A = \left.\frac{\partial^2 L}{\partial X^2}\right|_{X_0,Y_0}, \quad B = \left.\frac{\partial^2 L}{\partial Y^2}\right|_{X_0,Y_0}, \quad C = \left.\frac{\partial^2 L}{\partial X \partial Y}\right|_{X_0,Y_0}.$$

   (X_0, Y_0) is the most probable value for the joined posterior of parameters X and Y, and L is the logarithmic joined posterior probability function. In general, the iso-contours of Q at k ($k = Q(X,Y)$) trace an ellipse. Specify the following properties of this ellipse: the ellipse's centre; the magnitude of the axes and the orientation of the ellipse. How do these properties relate to the parameters' uncertainties and (lack of) correlation?

2. Open questions involving derivations

   (a) (**15 points**) Given a data set $\{x_i : 0 \le i < N\}$ of N values, where each value x_i is independent and drawn from a normal (Gaussian) distribution. The values have different (known) standard deviations, given by σ_i. Use Bayes' theorem (derived in question 1(a)) to prove that the most probable value μ_0 for the mean of the set is the weighted average of the data. Assume a flat prior for the mean μ.

→ **See next page for questions 2(b) and 3**

(b) (**12 points**) Assume a coin with unknown bias $H \in [0 \ldots 1]$ is thrown N times ($H = 0.5$ for a fair coin, $H = 1$ for a coin that always throws heads). The number of times heads comes up is R. Start from Bayes' theory (derived in question 1(a)) and derive the best estimate of the bias, H_0, and its variance σ_H^2, given R and N.

3. True/false questions – mark T for a true statement or F for a false statement on your exam paper: **1 point/question**

   (a) The Cauchy distribution is a specific instance of the more generic Student-T distribution.

   (b) Given two values that are from the same uniform distribution with zero mean, the sum of the squared values is chi-squared distributed.

   (c) A χ^2 test can be used to compare two binned distributions.

   (d) Compared to the linear correlation coefficient, Spearman's rank-order correlation coefficient is more robust against outliers.

   (e) The theory of "Ockham's razor" states that, in some situations, when independent random variables are added, their sum tends toward a normal distribution.

   (f) The Cauchy distribution has an undefined mean. (Was: "A set of samples that follow a Cauchy distribution has an undefined mean.")

   (g) For a uniform distribution, the cumulative distribution function and the probability distribution function are the same.

   (h) The mean of the Poisson distribution is equal to its variance.

   (i) Obtaining $\text{prob}(A \mid I)$ from $\text{prob}(A, B \mid I)$ requires marginalization over parameter A.

   (j) After calculating the posterior for a parameter for some set of data, the accuracy of the estimate can be increased by using the posterior as prior.