

Revision Cribsheet for 2021/22

As you know, the industrial action has had a severe impact upon the delivery of MAS2903 this year. For each chapter in the course notes, this cribsheet will tell you:

- (i) What is, and is not, examinable
- (ii) What you should be able to do, in the exam, based on the material that is examinable
- (iii) What you might want to memorise for the exam, and what you will not need to memorise

Chapter 1: Introduction

What is/is not examinable, as a result of the industrial action?

- Everything up to the end of Section 1.3 is examinable. That includes all material from page 1 through to the end of page 14. In lectures and/or asynchronous videos posted to Canvas, all boxes were filled in from page 1 through to the end of page 14
- All definitions, proofs and exercises appearing in Section 1.1, 1.2 and 1.3 are examinable
- Section 1.4 is non-examinable. We did not cover this due to the industrial action. Thus, no questions around *Sufficiency* will appear in your exam. You will not be required to define sufficiency, show that a statistics is sufficient, or find a sufficient statistic. You will not be required to use the Factorisation Theorem
- In short, you can disregard everything from page 15 through to page 21 of the lecture notes. Don't bother with Examples 1.8 through to Example 1.11

What should I know for the exam?

You should:

- Know the difference between the classical/frequentist/subjective interpretations of probability
- Be able to give examples of events whose probabilities can be found using the classical/frequentist/subjective interpretations of probability
- Be able to explain if the probabilities of events can be found using the classical/frequentist/subjective interpretations of probability
- Be able to explain the drawbacks of the classical/frequentist/subjective interpretations of probability
- Note: Much of this was bookwork, to be completed in your own time, by reading Section 1.2 of the notes. We will revisit this in class before the exam!

You should be able to:

- State the result for the *conditional probability* of one event given that another has occurred;
- Define a *partition* of the sample space
- State, and prove, the *Law of Total Probability*

- State, and prove, *Bayes Theorem*
- Apply Bayes Theorem to simple examples in probability
- Use Bayes Theorem to formulate diagnostic rules given some observed states/symptoms
- Find prior and posterior odds for partitioning events

You should:

- Know that the *likelihood function* for a parameter θ is simply the joint probability density function after observing data $\mathbf{x}$; assuming i.i.d. observations, this gives

$$f(\mathbf{x}|\theta) = \prod_{i=1}^n f_X(x_i|\theta)$$

- Be able to determine the likelihood function for commonly-used statistical models, or indeed *any* given probability density function for data $\mathbf{x}$

What would/would not be worth memorising for the exam?

Worth memorising:

- How to calculate probabilities using the classical and frequency interpretations
- Drawbacks of the frequency, classical and subjective approaches to probability
- The formula for finding conditional probabilities:

$$\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}$$

- The Law of Total Probability and its proof
- Bayes Theorem for conditional probability, and its proof:

$$\Pr(E_i|F) = \frac{\Pr(F|E_i)\Pr(E_i)}{\sum_{j=1}^n \Pr(F|E_j)\Pr(E_j)}$$

- The formula for finding prior and posterior odds of partitioning events
- The method for finding the likelihood function, given a set of i.i.d. observations

Not worth memorising:

- The condition of *Coherence*, as discussed under the subjective interpretation of probability at the bottom of page 4
- Anything relating to sufficiency in Section 1.4 of the notes, as all material in this section is non-examinable this year

Chapter 2: Bayes Theorem for Distributions

What is/is not examinable, as a result of the industrial action?

- Everything in Section 2.1 and Section 2.2 is examinable. This includes material from page 23 through to the end of Example 2.7 at the top of page 48
- As was the case for Chapter 1, the material relating to *Sufficiency* in Chapter 2 is non-examinable. This relates specially to Section 2.3 – everything from half-way down page 48 through to the end of page 51
- You can disregard Examples 2.8 and 2.9. In the exam, you will not be asked to determine a sufficient statistic, and you will not be asked to verify that the posterior given this sufficient statistic is identical to the posterior given the full sample; i.e. you will not be asked questions including the statement “Verify that $\pi(\theta|\mathbf{x}) = \pi(\theta|t)$ ”

What should I know for the exam?

You should be able to state Bayes’ Theorem for distributions, and you should know that

$$\text{posterior} \propto \text{prior} \times \text{likelihood}.$$

You should be familiar with the continuous Uniform distribution, the beta distribution and the gamma distribution, although you need *not* memorise their probability density functions, expectations and variances.

For a range of models $f(\mathbf{x}|\theta)$ for data $\mathbf{x}$, you should:

- Be able to obtain the posterior for θ given a prior $\pi(\theta)$. Examples include (but are not limited to):

$X_i \sim$	Prior for θ	Posterior
$Bin(n, \theta)$	Uniform	Beta
	Beta	Beta
$Po(\theta)$	Exp	Gamma
	Gamma	Gamma
$Exp(\theta)$	Gamma	Gamma
$N(\theta, 1/\tau)$	Normal	Normal

- Be able to work with specific (numerical) scenarios using observed data $\mathbf{x}$, as well as more general (algebraic) scenarios for $\mathbf{X} = X_1, X_2, \dots, X_n$
- Be able to find posterior distributions given unfamiliar distributions for data and prior beliefs

You should be able to describe how prior beliefs about a parameter of interest θ change in light of some observed data:

- Via prior and posterior numerical summaries
- Via plots/sketches of prior and posterior densities (make sure you can produce simple sketches of distributions and interpret plots of distributions)
- Via the computation of probabilities in the case where the prior/posterior are Normally distributed
- Be prepared to write short summaries, in plain English, relating to changes from the prior to the posterior. How has the average or mode changed? What about the variability or precision?

What would/would not be worth memorising for the exam?

Worth memorising:

- The main result of the course:

$$\text{posterior} \propto \text{prior} \times \text{likelihood}.$$

- The general form of the probability mass/density functions for some standard distributions – e.g. binomial, Poisson, exponential. You are not *required* to memorise these, as all important formulae will be provided on a formula sheet on the first page of the exam paper. However, being familiar with these will certainly help to speed up calculations and will help you to quickly spot the general form of priors and posteriors in the exam
- The derivation, and results, for Example 2.5
- The derivation, and results, for Example 2.6. You might be asked to derive the posterior, or you might just be asked a question that uses the results of this derivation. So make sure you're familiar with both the derivation and the results, specifically

$$\mu|\mathbf{x} \sim N(B, 1/D),$$

where

$$D = d + n\tau \quad \text{and} \quad B = \frac{db + n\tau\bar{x}}{D}.$$

- What posterior do you get when you combine a particular likelihood and prior? For example, Binomial (likelihood) + Beta (Prior) = Beta (Posterior)
- How to compute probabilities from the Normal distribution

Not worth memorising:

- Full details of probability mass/density functions, expectations, modes and variances of standard distributions. Although knowing the general form of the probability mass/density functions will certainly speed up calculations, and help you spot the form of the prior/posterior more quickly, full details will be given on the formula sheet in the exam
- The formulae for the gamma and beta functions, as given on page 25 in the notes – these will be provided on the formula sheet in the exam
- The likelihood modes for standard distributions – although these are often used in summary tables throughout chapter 2, using simple prior/posterior summaries will suffice
- Anything relating to sufficiency in Section 2.3 of the notes, as all material in this section is non-examinable this year

Chapter 3: Priors

What is/is not examinable, as a result of the industrial action?

- Everything in Section 3.1 and 3.2 is examinable. This includes all material from page 53 through to the end of page 67. We completed all parts of these sections during lectures
- Section 3.3 on *Parameter constraints* is non-examinable – we did not cover this as a result of the industrial action. This includes all material from page 68 through to the end of page 72
- *Almost* everything in the rest of Chapter 3 is examinable, including material in Section 3.4 and Section 3.5. We covered *Jeffreys priors* (Section 3.4) in an asynchronous video just before the Easter break, and *Asymptotic posteriors* (Section 3.5) will be covered in our first lecture back after the Easter break. I said “Almost”, because...
- ...the proof of the asymptotic posterior (pages 82-83) is non-examinable – not because of the industrial action, but because it has never been examinable. It is included in the notes for completeness

What should I know for the exam?

You should know how to obtain parameters for a prior distribution for θ given some suggested prior summaries for θ . For example:

- Parameters for the gamma distribution for θ given $E(\theta)$ and $Var(\theta)$
- Parameters for the beta distribution for θ given the mode for θ and a probability for θ

You should be able to describe the *trial roulette* and *bisection* methods of elicitation. You should also be able to:

- Interpret output from the *MATCH* elicitation tool
- Explain the processes of *feedback* and *refinement*
- Explain the role of *feedback percentiles*, and be able to transpose these into suitable questions to ask a non-statistician

You should be able to explain how we might operate Bayes Theorem with *vague prior knowledge*, and you should be able to obtain posterior distributions in circumstances when our prior knowledge about θ is indeed vague.

You should be able to explain *prior ignorance*; you should know how to construct ignorance priors for θ , including Jeffreys priors.

Given a model $f(\mathbf{x}|\theta)$, you should be able to obtain the asymptotic posterior distribution for θ .

What would/would not be worth memorising for the exam?

Worth memorising:

- How both the trial roulette and bisection methods of online elicitation work, and the advantages and disadvantages of both
- The meaning of substantial and vague prior knowledge, and the implications of both for the posterior:

Substantial prior knowledge : $\pi(\theta|\mathbf{x}) \sim \pi(\theta)$

Vague prior knowledge : $\pi(\theta|\mathbf{x}) \sim f(\mathbf{x}|\theta)$

- The conjugate priors to use in conjunction with some standard models for data, in order to obtain posteriors assuming vague prior knowledge. For example:

$X_i \sim$	Conjugate prior for θ	Posterior
$Po(\theta)$	Gamma	Gamma
$Exp(\theta)$	Gamma	Gamma
$N(\theta, 1/\tau)$	Normal	Normal

- The formula for Jeffreys prior:

$$\pi(\theta) \propto \sqrt{I(\theta)},$$

where

$$I(\theta) = E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right].$$

- The formula for the asymptotic posterior:

$$\theta|\mathbf{x} \sim N\left(\hat{\theta}, J(\hat{\theta})^{-1}\right),$$

approximately, as $n \rightarrow \infty$, where

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta)$$

and $\hat{\theta}$ is the maximum likelihood estimate of θ .

Not worth memorising:

- Any of the formulae related to expectations, modes and variances of the standard distributions – although we use these in this chapter (e.g. Example 3.1, Example 3.2), you will be given these on the formula sheet in the exam
- Any R code, as used in some examples in this chapter (e.g. Example 3.2). However, it is worth familiarising yourself with such code as you could be presented with code in the exam and be asked to interpret it
- Any definitions, theorems, proofs or examples in Section 3.3, as this section is non-examinable due to the effects of the industrial action
- The proof of the asymptotic posterior distribution, shown on pages 82/83 of the course notes. This has never been examinable in this module

Chapter 4: Bayesian Inference

What is/is not examinable, as a result of the industrial action?

I am writing this revision cribsheet just before we return from the Easter vacation, so it's difficult for me to know exactly what we will have time to cover in Chapter 4 in the remaining lecture time for semester 2. Obviously the industrial action before Easter has set us back somewhat. The following is what I anticipate will be the case.

- Everything in Section 4.1 is examinable. This includes differences in interpretation of Bayesian and frequentist confidence intervals, and the computation of Bayesian confidence intervals in simple cases (e.g. Examples 4.1, 4.2 and 4.3)
- Only some parts of Section 4.2 on *Prediction* will be examinable, including:
 - The construction of the predictive distribution in simple cases (e.g. Examples 4.4 and 4.5)
 - The construction of prediction intervals (e.g. as illustrated in Examples 4.4 and 4.5)
- Material that will be non-examinable from chapter 4 includes:
 - Differences between “correct” and “naive” predictive probability functions (e.g. as illustrated in Table 4.3 and discussed on page 96)
 - The alternative method for finding the predictive distribution (using “Candidate’s formula”), as discussed at the bottom of page 99/top of page 100
 - The construction of predictive distributions in more complicated cases (e.g. Example 4.6)

What should I know for the exam?

You should:

- Understand the difference between a *point estimate* and an *interval estimate*
- Be able to define a *frequentist confidence interval*, a *Bayesian confidence interval* and a *highest density interval*, and you should understand the differences in interpretation between these intervals
- At this point, we compare Bayesian and frequentist inference more generally. Therefore, you should be able to describe the benefits of a Bayesian analysis over a frequentist analysis
- Be able to determine a highest density interval for θ in the following simple situations:
 - When the posterior for θ is a beta distribution and the lower/upper bounds of the interval are clearly 0/1
 - When the posterior for θ is symmetric, e.g. a Normal distribution

You should be able to:

- Interpret simple functions in R to obtain a highest density interval for any unimodal posterior distribution
- Describe, through words and algebraically, how highest density intervals would be constructed for any unimodal posterior distribution

You should be able to use the following result to obtain the predictive probability density function for a future observation y :

$$f(y|\mathbf{x}) = \int_{\Theta} f(y|\theta) \pi(\theta|\mathbf{x}) d\theta.$$

You should be able to obtain prediction intervals in simple examples for discrete θ , or when the predictive distribution is of known form.

What would/would not be worth memorising for the exam?

Worth memorising:

- The differences in interpretation between Bayesian and frequentist confidence intervals, and the definition of a Bayesian *highest density interval*
- The critical values from the standard Normal distribution for the construction of symmetric confidence intervals when the posterior is Normal – for example, $z = 1.96$ and $z = 2.58$ for a 95% and a 99% confidence interval (respectively)
- The formula for the predictive probability density function:

$$f(y|\mathbf{x}) = \int_{\Theta} f(y|\theta) \pi(\theta|\mathbf{x}) d\theta.$$

Not worth memorising:

- The formulae associated with the *Beta-Binomial* distribution, as provided in Definition 4.1 - if you have to use this, it will be given in the formula sheet in the exam
- The alternative formula for the predictive probability density function:

$$f(y|\mathbf{x}) = \frac{f(y|\theta)\pi(\theta|\mathbf{x})}{\pi(\theta|\mathbf{x}, y)}$$

- The formulae for the *Inverse-Beta* distribution, as given in Definition 4.2. You will not have to use this in the exam

Chapter 5

- Questions 4, 10, 11, 21, 30 and 42 were used in the assignments for the module and, as such, as all examinable
- Given the industrial action, and parts of the course that have been deemed non-examinable as a result, questions 7, 8, 9, 14, 20, 28, 31, 32, 34, 40 and 41 can be disregarded when preparing for the exam
- All other questions should be useful when preparing for the exam