Questions 1.15, 1.18, 2.14, 2.15, 2.17, 3.10–3.18 inc., 4.16, 4.17, 4.18 and all questions on Chapters 5–9 are not relevant to the 2010–2011 course.

F.1 Further Exercises on Chapter 1

1. Suppose a population consists of 100 individuals, of whom 10 per cent have high blood pressure. What is the probability that at most two out of eight individuals chosen at random from this population have high blood pressure?

2. Consider a collection of families in which the sexes of successive children are independent, both sexes being equally likely. Show that in families of exactly three children, the event that the family has children of both sexes is independent of the event that there is at most one girl, but that this independence does not hold among families of of exactly two children or among families of exactly four children.

3. It is known that 5% of all men and 0.25% of all women are colour-blind. A person chosen by chance suffers from colour-blindness. What is the probability that this is a man? (It is assumed that there is an equal number of men and women).

4. At a factory where bolts are produced, machines A, B, C produce respectively 25%, 35% and 40% of all bolts. Of the output of these machines, defects constitute 5%, 4% and 2% respectively. A bolt selected at random from the production turned out to be defective. What is the probability that it was produced by machine A? machine B? machine C?

5. Suppose a red die and a blue die are tossed. Let x be the sum of the number showing on the red die and three times the number showing on the blue die. Find the density function and the distribution function of x,

6. Show that the probability density function

$$p(x) = \binom{n+x-1}{x} \pi^n (1-\pi)^x$$

of the negative binomial distribution $\text{NB}(n, \pi)$ of index n and parameter π (as defined in Section A.13, page 284 of the book) tends to that of the Poisson distribution $\text{P}(\lambda)$ of mean λ as $n \to \infty$ while $\pi \to 1$ in such a way that $n(1-\pi) = \lambda$ where λ is constant.

7. Suppose that i and j have independent binomial distributions with the

same parameter π but different parameters m and n, so that $i \sim \mathrm{B}(m, \pi)$ and $j \sim \mathrm{B}(n, \pi)$, and let $k = i + j$.

(a) Show that k has a binomial distribution which also has parameter π, but has index of $m + n$.

(b) Find the distribution of i conditional on the value of k.

8. Find expressions for (a) the density $u = x^3$ and (b) the density of $v = x^4$ in terms of that of x, in both the continuous and discrete cases and find the resulting densities in the case where x has a standard normal density. What is the most important difference between the two examples?

9. Suppose that x_1, x_2 and x_3 are independent and all have the same continuous distribution with density $p(x)$ and distribution function $F(x)$. Find the distribution function of

$$x_{\text{med}} = \text{median}\,\{x_1, x_2, x_3\}$$

in terms of $F(x)$, and thus find an expression for the density function of x_{med}.

10. Given $p(x, y) = xy \exp[-\frac{1}{2}(x^2+y^2)]$ where $x > 0$ and $y > 0$ and $p(x, y) = 0$ otherwise, calculate (a) $\mathsf{P}(x < 1)$, (b) $\mathsf{P}(x < 1, y < 1)$.

11. Show that $\mathsf{E}\, g(\widetilde{x})h(\widetilde{y}) = (\mathsf{E}g(\widetilde{x}))(\mathsf{E}h(\widetilde{y}))$ for arbitrary functions g and h if and only if $\widetilde{x}$ and $\widetilde{y}$ are independent.

12. Suppose that the random variable k has a logarithmic series distribution with parameter θ (where $0 < \theta < 1$), so that

$$p(k) = \frac{\alpha\theta^k}{k} \qquad (k = 1, 2, \dots)$$

where $\alpha = -[\log(1-\theta)]^{-1}$ and thus the probabilities are the terms in the series expansion of $-\alpha\log(1-\theta)$. Find the mean and variance of k.

13. A random variable x is said to have a Weibull distribution if there are values of the parameters c (> 0) and α (> 0) such that

$$p(x) = c\alpha^{-1}\{x/\alpha\}^{c-1}\exp[-\{x/\alpha\}^c]$$

for $x > 0$ and otherwise $p(x) = 0$. Find a function y of x such that y has an exponential distribution (see the last subsection of Section 2.5 or Section Section A.4 in the book) and hence find the mean and variance of x.

14. The *skewness* of a random variable x is defined as $\gamma_1 = \mu_3/(\mu_2)^{3/2}$ where

$$\mu_n = \mathsf{E}(x - \mathsf{E}x)^n$$

(but note that some authors work in terms of $\beta_1 = \gamma_1^2$). Find the skewness of a random variable X with a Poisson distribution $\mathrm{P}(\mu)$ of parameter μ.

15. Show that for any random variable X and any constants $t > 0$ and $c > 0$

$$\mathsf{E}(X + c)^2 \geqslant (t + c)^2 \mathsf{P}(X > t)$$

and deduce from the Parallel Axes Theorem or otherwise that if X has mean 0 and variance σ^2 then

$$\mathsf{P}(X > t) \leqslant \frac{\sigma^2 + c^2}{(t + c)^2}.$$

By minimizing the right-hand side as a function of c show that

$$\mathsf{P}(X > t) \leqslant \frac{\sigma^2}{\sigma^2 + t^2}.$$

16. Suppose that the point (x, y) is uniformly distributed over the interior of the circle $x^2 + y^2 = 1$. Show that x and y are uncorrelated but that they are *not* independent.

17. Let $\boldsymbol{x} = (x_1, x_2, x_3)$ be a continuous random vector having joint density

$$p(x_1, x_2, x_3) = \begin{cases} 6\exp(-x_1 - x_2 - x_3) & (0 < x_1 < x_2 < x_3) \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the marginal density function of x_2, (b) the conditional density of (x_1, x_3) given x_2, (c) the marginal (joint) density of (x_1, x_2), and (d) the conditional density of x_3 given (x_1, x_2), and (e) the conditional density of x_1 given x_2.

18. Suppose that $\widetilde{\lambda}$ has a (negative) exponential distribution with parameter 1 (see Section A.4 in the book), so that its mean and variance are both equal to unity. Suppose further that conditional on $\widetilde{\lambda} = \lambda$ the value of k has a Poisson distribution of mean λ, so that $\mathsf{E}_{\widetilde{k}|\widetilde{\lambda}}\left(\widetilde{k}|\widetilde{\lambda}\right) = \widetilde{\lambda}$ and $\mathcal{V}_{\widetilde{k}|\widetilde{\lambda}}\left(\widetilde{k}|\widetilde{\lambda}\right) = \widetilde{\lambda}$.

(a) Show by integrating the joint density $p(k, \lambda) = p(\lambda)p(k|\lambda)$ over λ that the unconditional distribution of $\widetilde{k}$ is geometric with $\pi = \frac{1}{2}$ (see Section A.13 in the book), so that $p(k) = 1/2^{k+1}$.

(b) Verify that the formula

$$\mathcal{V}\widetilde{k} = \mathsf{E}_{\widetilde{\lambda}}\mathcal{V}_{\widetilde{k}|\widetilde{\lambda}}\left(\widetilde{k}|\widetilde{\lambda}\right) + \mathcal{V}_{\widetilde{\lambda}}\mathsf{E}_{\widetilde{k}|\widetilde{\lambda}}\left(\widetilde{k}|\widetilde{\lambda}\right)$$

derived in Section 1.5 holds in this case.

F.2 Further Exercises on Chapter 2

1. Suppose that $k \sim \mathrm{P}(\lambda)$. Find the standardized likelihood as a function of λ for given k. Which of the distributions listed in Appendix A does this represent?

2. Suppose we are given the following 11 observations from a normal distribution:

$$148,\ 154,\ 158,\ 160,\ 161,\ 162,\ 166,\ 170,\ 182,\ 195,\ 236$$

and we are told that the standard deviation $\sqrt{\phi} = 25$. Find a 95% HDR for the posterior distribution of the mean assuming the usual reference prior.

3. With the same data as in the previous question, what is the predictive distribution for a possible future observation x?

4. Show that if a random sample of size $n = \phi^2$ is taken from an $\mathrm{N}(\theta, \phi)$ distribution where θ has a prior distribution which also has a variance of ϕ, then the posterior distribution of θ cannot also have variance ϕ.

5. Your prior beliefs about a quantity θ are such that

$$p(\theta) = \begin{cases} 1 & (\theta \geqslant 0) \\ 0 & (\theta < 0). \end{cases}$$

A random sample of size 16 is taken from an $\mathrm{N}(\theta, 1)$ distribution and the mean of the observations is observed to be 0.45. Find a 90% HDR for θ.

6. Suppose that you have prior beliefs about an unknown quantity θ which can be approximated by an $\mathrm{N}(\lambda, \phi)$ distribution, while my beliefs can be approximated by an $\mathrm{N}(\mu, \phi)$ distribution (so that the variances are the same although the means are different). Suppose further that the reasons that have led us to these conclusions do not overlap with one another. What distribution should represent our beliefs about θ when we take into account all the information available to both of us?

7. Suppose that in the situation described in Section 2.4 $\lambda_\alpha = 1.96$ and the prior density is constant within I_α and that prior to taking the sample no value of θ outside I_α is more than twice as probable as any value of θ inside it. How far can the true posterior density differ from the normal density within I_α? What difference does it make if λ_α is increased to 3.29?

8. Which of the following can be expressed in data-translated form:

(a) Any likelihood of the form

$$l(\sigma|\boldsymbol{y}) \propto h\left(\frac{s(\boldsymbol{y})}{\sigma}\right);$$

(b) The likelihood

$$l(\xi_0|\boldsymbol{x}) \propto \prod p(x_i \,|\, \xi_0, \alpha, c)$$

where $p(x \,|\, \xi_0, \alpha, c)$ is the Weibull density

$$p(x \,|\, \xi_0, \alpha, c) = c\alpha^{-1}\{(x - \xi_0)/\alpha\}^{c-1} \exp[-\{(x - \xi_0)/\alpha\}^c] \qquad (\xi_0 < x);$$

(b) The likelihood

$$l(\alpha|\boldsymbol{x}) \propto \prod p(x_i \,|\, 0, \alpha, c)$$

where $p(x \,|\, \xi_0, \alpha, c)$ is the Weibull density as above with $\xi_0 = 0$.

9. Suppose you are interested in investigating how variable is the performance of schoolchildren on a new English test, and that you begin by trying this test out on children in 15 similar schools. It turns out that the average standard deviation is about 20 marks. You then want to try the test on a sixteenth school, which is fairly similar to those you have already investigated, and you reckon that the data on the other schools gives you a prior for the variance in this new school which has a mean of 120 and is worth nine direct observations on the school. What is the posterior distribution for the variance if you observe a sample of size 50 from the school of which the standard deviation is 28.9? Give an interval in which the variance lies with 95% posterior probability.

10. The following are malt-extract values on malts made from Kindred barley grown at various locations in the Mississippi Valley Barley Nurseries during 1948 in percent of dry basis:

77.7, 76.0, 76.9, 74.6, 74.7, 76.5, 74.2, 75.4, 76.0, 76.0, 73.9, 77.4, 76.6, 77.3.

Give 99% HDRs for the mean and variance of the population from which they come.

11. Show that if $x_j \sim \text{Be}(\alpha, \beta)$ (cf. Section A.10 in the book) for $i = 1, 2, \dots,$ n, then $(\prod x_j, \prod(1 - x_j))$ is sufficient for (α, β) given $\boldsymbol{x}$. Show further that if α is known, then $\prod(1 - x_j)$ is sufficient for β and that if β is known then $\prod x_j$ is sufficient for α.

12. Suppose that $x_1, \dots, x_n$ are independent, identically distributed random variables with discrete probability density

$$p(x \,|\, \pi, \theta) = (1 - \pi)\pi^{x-\theta}, \qquad (x = \theta, \theta + 1, \theta + 2, \dots),$$

where θ and π are unknown parameters and $0 < \pi < 1$. Show that the vector $(\min x_j, \sum x_j)$ is sufficient for (θ, π). Show further that if π is known, then $\min x_j$ is sufficient for θ, and that if θ is known, then $\sum x_j$ is sufficient for π.

13. Suppose that x has a Poisson distribution of mean $1/\lambda^2$, so that the likelihood takes the form

$$l(\lambda|x) = \lambda^{-2x} \exp(-1/\lambda^2).$$

Find a family of conjugate priors for λ.

14. Show that if y and z are independent standard normal variates then

$$\frac{y+z}{y-z} \sim \mathrm{C}(0,1).$$

(*Hint:* Note that a $\mathrm{C}(0,1)$ distribution is the same thing as a t distribution on 1 degree of freedom; see Section A.21 in the book.)

15. Suppose that $x \sim \mathrm{G}(\alpha, \beta)$ has a two-parameter gamma distribution (see Section A.4 in the book) and that $y-x$ has the same distribution (this situation arises when x is the lifetime of some component which has to be replaced and y is the time when it is replaced for a second time). Show that the vector (x, y) has a distribution in the two-parameter exponential family.

16. Suppose that the results of a certain test are known, on the basis of general theory, to be normally distributed about the same mean μ with the same variance ϕ, neither of which is known. Suppose further that your prior beliefs about (μ, ϕ) can be represented by a normal/chi-squared distribution with

$$\nu_0 = 5, \qquad S_0 = 250, \qquad n_0 = 2, \qquad \theta_0 = 115.$$

Now suppose that 50 observations are obtained from the population with mean 77 and sample variance $s^2 = 20$. Find the posterior distribution of (μ, ϕ). Compare 50% prior and posterior HDRs for μ.

17. Suppose that your prior for θ is a $\frac{3}{4} : \frac{1}{4}$ mixture of $\mathrm{N}(0,1)$ and $\mathrm{N}(0.5,1)$ and that a single observation $x \sim \mathrm{N}(\theta, 1)$ turns out to equal 1.5. What is your posterior probability that $\theta > 1$?

18. An experimental station has had experience with growing wheat which leads it to believe that the yield per plot is more or less normally distributed with mean 200 and standard deviation 15. The station then wished to investigate the effect of a growth hormone on the yield per plot. In the absence of any other information, the prior distribution for the variance on the plots might

be taken to have mean 200 and standard deviation 90. As for the mean, it is expected to be about 230, and this information is thought to be worth about 20 observations. Twelve plots treated with the hormone gave the following yields:

222, 234, 156, 287, 190, 255, 307, 101, 133, 251, 177, 225,

Find the posterior distributions of the mean and variance.

F.3 Further Exercises on Chapter 3

1. Give the form of the beta-binomial predictive distribution for the number of successes in the next m trials after you have observed n trials in which there were x successes assuming that you used the standard Haldane prior $\text{Be}(0, 0)$.

2. Find a suitable interval of 95% posterior probability to quote in a case where your posterior distribution for an unknown parameter π is $\text{Be}(18, 13)$ and compare this with interval with similar intervals for the cases of $\text{Be}(18.5, 13.5)$ and $\text{Be}(19, 14)$ posteriors. Comment on the relevance of the results to the choice of a reference prior for the binomial distribution.

3. Suppose that your prior beliefs about the probability π of success in Bernoulli trials have mean $1/4$ and variance $1/48$. Give a 90% HDR for π given that you have observed six successes in 10 trials.

4. Suppose that you have a prior distribution for the probability π of success in a certain kind of gambling game which has mean 0.7, and that you regard your prior information as equivalent to 16 trials. You then play the game 30 times and win 18 times. What is your posterior distribution for π?

5. Suppose that you are interested in the proportion of females in a certain organisation and that as a first step in your investigation you intend to find out the sex of the first 13 members on the membership list. Before doing so, you have prior beliefs which you regard as equivalent to 30% of this data, and your prior beliefs suggest that a 50% of the membership is female.

Suggest a suitable prior distribution and find its standard deviation.

Suppose that 4 of the first 13 members turn out to be female; find your posterior distribution and give a 50% posterior HDR for this distribution.

Find the mean, median and mode of the posterior distribution.

Would it surprise you to learn that in fact 149 of the total number of 551 members are female?

6. Suppose that the distribution of a random variable x is such that the standard deviation is always equal to the mean μ. Show that $\log x$ has a variance which is approximately constant. What does this suggest as a reference prior for μ?

7. The following data (quoted by Smith, 1969, Section 20.12 from an article by L F Richardson) give the observed numbers of outbreaks of war per year for the 432 years from 1500 to 1931 A.D.

Number of wars:	0	1	2	3	4	5 and more
Number of years:	223	142	48	15	4	0.

Give an interval in which the mean number λ of outbreaks of war per year lies with 95% probability.

8. Recalculate your answer to the preceding question assuming that you had a prior distribution for λ of mean 0.5 and standard deviation 0.16.

9. We say that x has a Weibull distribution with parameters μ, σ and c it it has density

$$\left[\frac{c}{\sigma}\right]\left[\frac{x-\mu}{\sigma}\right]^{c-1}\exp\left[-\left[\frac{x-\mu}{\sigma}\right]^{c}\right] \qquad (\mu \leqslant x < \infty)$$

Suppose that $\mu = 0$ and that c is known. What prior for σ is suggested by Jeffreys' rule?

10. Suppose that the probability density function of an observation x depends on two parameters θ and ϕ in such a way that it can be expressed as a product of a function of θ and x and a function of ϕ and x, and you find the prior given by the two-dimensional version of Jeffreys' rule. Show that this prior gives the pair (θ, ϕ) the distribution which results from giving θ the Jeffreys' prior obtained by regarding ϕ as constant, and ϕ the Jeffreys' prior obtained by regarding θ as constant, independently of one another.

11. A generalized Pareto distribution due to M Ljubo is defined by Johnson *et al.* (1994–1995) by the distribution function

$$F(x) = 1 - \left(\frac{\xi+\alpha}{x+\alpha}\right)^{\gamma} \mathrm{e}^{-\beta(x-\xi)} \qquad (x > \xi)$$

(and $F(x) = 0$ for $x \leqslant \xi$). Consider the case where $\beta = 0$ and ξ and γ are known. Use Jeffreys' rule to find a suitable reference prior for α.

12. Consider a uniform distribution on the interval (α, β), where the values of α and β are unknown, and suppose that the joint distribution of α and β is a bilateral bivariate Pareto distribution with $\gamma = 2$. How large a random sample must be taken from the uniform distribution in order that the coefficient of variation (i.e. the standard deviation divided by the mean) of the length $\beta - \alpha$ of the interval should be reduced to 0.01 or less?

13. Suppose that observations x_1, x_2, $\dots$, x_n are available from a density

$$p(x|\theta) \propto \exp(-\theta/x) \qquad (0 < x < \theta).$$

Explain how you would make inferences about the parameter θ using a conjugate prior.

14. Give an approximation to the median of the posterior distribution for the number of tramcars given than you have observed n tramcars numbered x_1, x_2, $\dots$, x_n.

15. Test the assertion of Benford's Law by taking a table of physical constants (e.g. the one in Abramowitz and Stegun, 1964 or 1965, Table 2.3) or tables in other standard reference works and seeing how well it fits.

16. We sometimes investigate distributions on the vertices of a regular solid, for example a cube or a dodecahedron. Find a Haar prior for a distribution on the vertices of such a solid.

17. Suppose that the prior distribution $p(\mu, c)$ for the parameters μ and c of a distribution with a density of the form

$$p(x \mid \mu, c) = (\text{const}) \times c^{-1} \exp\left(-\sum |x-\mu|^3/c^3\right)$$

is uniform in μ and c, and that the four observations

$$x_1 = 1, \qquad x_2 = 1, \qquad x_3 = 2, \qquad x_4 = 3,$$

are available from this distribution. Calculated the value of the posterior density $p(\mu, c)$ (ignoring the constant) to one decimal place for $\mu = 1$, 1.5, 2, 2.5, 3 and $c = 1.2$, 1.4, 1.6, 1.8, 2.0.

Use Simpson's rule

$$\int_a^b f(t)\,\mathrm{d}t = \frac{(b-a)}{3n}\{f(t_0) + 4f(t_1) + 2f(t_2) + 4f(t_3) + \dots + f(t_n)\}$$

to approximate the posterior density for μ.

Go on to find an approximation to the posterior probability that

$$1.75 < \mu < 2.25.$$

18. Fisher (1925b, Chapter 9, Example 47) quotes a genetic example which results in individuals of four types with probabilities $(2+\theta)/4$, $(1-\theta)/4$, $(1-\theta)/4$ and $\theta/4$ respectively, where θ is an unknown parameter between 0 and 1. Observations on n individuals which were independently of these four types showed that there were a, b, c, and d respectively. Find the maximum likelihood estimator and the Fisher information.

Show that if $a = 1997$, $b = 906$, $c = 904$ and $d = 32$ (so that $n = 3839$), then the value of the maximum likelihood estimator is 0.0357.

Show that the method of scoring starting from a simple inefficient estimator proposed by Fisher, namely

$$t = \{a + d - (b + c)\}/n$$

gets to within 0.001 of the maximum likelihood estimator in two iterations.

F.4 Further Exercises on Chapter 4

1. Show that if the prior probability π_0 of a hypothesis is close to zero, then the posterior probability p_0 satisfies $p_0 \cong 1 - \pi_0 B^{-1}$ and more exactly $p_0 \cong 1 - \pi_0 B^{-1} - \pi_0 \left(B^{-1} - B^{-2}\right)$.

2. Consider the data in Section 2.2 on the age of Ennerdale granophyre. Using the prior based on the K/Ar method and the posterior derived in the book, find the prior and posterior odds and the Bayes ratio for the hypothesis that the age is greater than 400 million years.

3. A manufacturing process is adjusted so that the mean of a certain dimension of the manufactured part is 20 cm. A sample of ten parts is checked as to this dimension, with the following results

$$19.66,\ 20.04,\ 19.96,\ 19.92,\ 20.04,\ 19.66,\ 19.88,\ 19.82,\ 20.10,\ 19.72.$$

Find the P-value for the hypothesis that the mean value has dropped significantly below the desired value of 20 cm.

4. Compare the observed fraction of zeroes in the first 250 digits in Table 7.1 of Neave (1978) with the expected number and use Lindley's method to see whether this data provides evidence that these digits are not truly random.

5. With the data due to von Bortkiewits (1898) quoted in question 7 of Chapter 3, would it be appropriate to reject the hypothesis that on average one man is killed per cavalry corps per year? [Use Lindley's method.]

6. Suppose that the standard test statistic $z = (\overline{x} - \theta_0)/\sqrt{(\phi/n)}$ takes the value $z = 3$ and that the sample size is $n = 64$. How close to θ_0 does a value of θ have to be for the value of the normal likelihood function at $\overline{x}$ to be within 5% of its value at $\theta = \theta_0$?

7. Show that in the situation described in Section 4.5 an observation z of order $n^{1/4-\varepsilon}$ will, for large n, result in a posterior probability very near unity for the null hypothesis.

8. Suppose that $x_1, x_2, \dots, x_n \sim \mathrm{N}(\theta_0, \phi)$ where ϕ is known, and take $\phi = 1$ and $n = 16$. Find a value of ε such that over the interval $(\theta_0 - \varepsilon, \theta_0 + \varepsilon)$ the likelihood given a value $\overline{x}$ which is $2\sqrt{(\phi/n)}$ from θ_0 does not vary by more than $2\frac{1}{2}\%$.

9. At the beginning of Section 4.5, we saw that under the alternative hypothesis

that $\theta \sim \mathrm{N}(\theta_0, \psi)$ the predictive density for $\overline{x}$ was $\mathrm{N}(\theta_0, \psi + \phi/n)$, so that

$$p_1(\overline{x}) = \{2\pi(\psi + \phi/n)\}^{-\frac{1}{2}} \exp[-\tfrac{1}{2}(\overline{x} - \theta_0)^2/(\psi + \phi/n)]$$

Find the maximum of this density considered as a function of n on the assumption that $z \geqslant 1$. Hence find an upper bound for the Bayes factor under this assumption.

10. Add a row to the table in the subsection 'Numerical examples' of Section 4.5 for 2-tailed P-value 0.02.

11. Mendel (1865) reported finding 2001 green seeds to 6022 yellow in an experiment in which his theory predicted a ratio of $1:3$. Use the method employed for Weldon's dice data in Section 4.5 to test whether his theory is confirmed by the data. [However, Fisher (1936) cast some doubt on the genuineness of the data.]

12. Find a value for a Bayes factor when testing $\mathrm{H}_0 : \lambda = \lambda_0$ against the alternative hypothesis $\mathrm{H}_1 : \theta \neq \theta_0$ given a single observation $x \sim \mathrm{P}(\lambda)$. Use a gamma prior with mean λ_0 and variance σ^2 under the alternative hypothesis.

13. Add a row to the table in the subsection 'A bound which does not depend on the prior distribution' of Section 4.5 for 2-tailed P-value 0.02.

14. Suppose that x has a Poisson distribution $\mathrm{P}(\lambda)$ of mean λ, and that it is desired to test $\mathrm{H}_0 : \lambda = \lambda_0$ against the alternative hypothesis $\mathrm{H}_1 : \theta \neq \theta_0$.

(a) Find lower bounds on the posterior probability of H_0 and on the Bayes factor for H_0 versus H_1, bounds which are valid for any $\rho_1(\theta)$.

(b) If $\lambda_0 = 2$ and $x = 6$ is observed, calculate the (two-tailed) P-value and the lower bound on the posterior probability when the prior probability π_0 of the null hypothesis is $\frac{1}{2}$.

15. Twenty observations from a normal distribution of mean θ and variance ϕ are available, of which the sample mean is 1.15 and the sample variance is 1.5. Compare the Bayes factors in favour of the null hypothesis that $\theta = \theta_0$ assuming (a) that ϕ is unknown and (b) that it is known that $\phi = 1$.

16. Rewrite the argument in Jeffreys (1961, Section 5.1) leading to

$$B \approx \sqrt{\left(\frac{\pi\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu-1)/2}$$

in your own words.

17. Suppose that in testing a point null hypothesis you find a value of the usual Student's t statistic of 3.9 on 11 degrees of freedom. Would the methodology of Section 4.6 require you to "think again"?

18. Show that for large samples the Doogian philosophy requires that the sample size be of order $1/P^2$ if you are not to have to "think again".

F.5 Further Exercises on Chapter 5

1. Smith (1969, Table 21.8) quotes the following data on haemoglobin concentration in ten animals before and after the administration of a drug:

Animal	1	2	3	4	5	6	7	8	9	10
Before	151	147	158	141	155	160	153	140	152	145
After	148	145	155	136	152	159	150	137	148	142

Investigate the mean discrepancy θ between their results and in particular give an interval in which you are 90% sure that it lies.

2. With the same data as in the previous question, test the hypothesis that there is no discrepancy between the two analysts.

3. Suppose that you have grounds for believing that observations x_i, y_i for $i = 1, 2, \ldots, n$ are such that $x_i \sim \mathrm{N}(\theta, \phi_i)$ while $y_i \sim \mathrm{N}(\theta, 2\phi_i)$, but that you are not prepared to assume that the ϕ_i are equal. What statistic would you expect to base inferences about θ on?

4. How much difference would it make to the analysis of the data in Section 5.1 on rat diet if we took $\omega = (\sqrt{\phi} + \sqrt{\psi})^2$ instead of $\omega = \phi + \psi$?

5. The values of oxygen consumption (in mm^3 per hour per gram weight) were measured for ten fish in rapid streams and for ten in slow-moving water. The values found are reported by Smith (1969, Section 22.5, Problem 1) as:

In rapid streams	117	117	122	116	135	144	106	114	99	108
In slow water	90	58	82	82	92	75	90	142	85	107

Investigate the mean discrepancy θ between the mean consumption and in particular give an interval in which you are 90% sure that it lies

(a) assuming that it is known from past experience that the standard deviation of both sets of observations is 20, and

(b) assuming simply that it is known that the standard deviations of the two sets of observations are equal.

6. A random sample $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ is available from an $\mathrm{N}(\lambda, \phi)$ distribution and a second independent random sample $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ is available from an $\mathrm{N}(\mu, 3\phi)$ distribution. Obtain, under the usual assumptions, the posterior distributions of $\lambda - \mu$ and of ϕ.

7. In situation described at the start of Section 5.2, what is the posterior distribution of the variance? Give a 90% HDR for the variance in the example on the weight growth of rats.

8. The following table gives the values of the cephalic index found in two random samples of skulls, one consisting of fifteen and the other of thirteen individuals:

Sample I:	74.1	77.7	74.4	74.0	73.8	79.3	75.8	
	82.8	72.2	75.2	78.2	77.1	78.2	76.3	76.8
Sample II:	70.8	74.9	74.2	70.4	69.2	72.2	76.8	
	72.4	77.4	78.1	72.8	74.3	74.7		

Investigate the difference θ between these lengths without making any particular assumptions about the variances of the two populations, and in particular give an interval in which you are 90% sure that it lies.

9. Show that if $m = n$ is large then f_1 in Patil's approximation is approximately equal to unity while f_2 is approximately $n^{-1}(1 - \frac{1}{2}\sin^2 2\theta)$.

10. Suppose that T_x, T_y and θ are defined as in Section 5.3 and that

$$T = T_x \sin\theta - T_y \cos\theta, \qquad U = T_x \cos\theta + T_y \sin\theta$$

Show that if $\gamma = \lambda - \mu$ then

$$U = \frac{\gamma - (\overline{x} + \overline{y})}{\sqrt{(s_x^2/m + s_y^2/n)}} \qquad \text{and} \qquad \tan\theta = \frac{s_x/\sqrt{m}}{s_y/\sqrt{n}}$$

Show further that $U \sim \mathrm{BF}(\nu_x, \nu_y, \theta + \frac{1}{2}\pi)$. Can this fact be used to test the hypothesis that $\lambda + \mu = 0$?

11. Show that if $x \sim \mathrm{z}_{\nu_1,\nu_2}$ then

$$\frac{\nu_1 \mathrm{e}^{2x}}{\nu_2 + \nu_1 \mathrm{e}^{2x}} \sim \mathrm{Be}(\tfrac{1}{2}\nu_1, \tfrac{1}{2}\nu_2).$$

12. P G Hoel, *An Introduction to Mathematical Statistics* (5th edn), New York: Wiley 1984, Chapter 10, Section 5.2, quotes the following data on the yield of corn in bushels per plot on 20 experimental plots of ground, half of which were treated with phosphorus as a fertilizer:

Treated	6.2	5.7	6.5	6.0	6.3	5.8	5.7	6.0	6.0	5.8
Untreated	5.6	5.9	5.6	5.7	5.8	5.7	6.0	5.5	5.7	5.5

Find the sample variances for the treated plots and for the untreated plots, and give an interval in which you are 95% sure that the ratio of the population variances lies.

13. Measurement errors when using two different instruments are more or less symmetrically distributed and are believed to be reasonably well approximated by a normal distribution. Twenty measurements with each show a sample standard deviation twice as large with one instrument as with the other. Give an interval in which you are 90% sure that the ratio of the true standard deviations lies.

14. Repeat the analysis of Di Raimondo's data in Section 5.6 on the effects of penicillin of mice, this time assuming that you have prior knowledge worth about twelve observations in each case suggesting that the mean chance of survival is about a quarter with the standard injection but about three-quarters with the penicillin injection.

15. Lindley (1965, Sections 7.4 and 7.6) quotes a genetical example involving two genes which can be summarized as follows:

	B	$\overline{B}$	Total
A	6	13	19
$\overline{A}$	16	61	61
Total	22	74	96

Find the posterior probability that the log odds-ratio is positive and compare it with the comparable probability found by using the inverse root-sine transformation.

16. Show that if $\pi \cong \rho$ then the log odds-ratio is such that

$$\Lambda - \Lambda' \cong \frac{\alpha(1-2\alpha)}{\pi(1-\pi)} + \frac{\alpha^2}{\pi^2(1-\pi)^2}$$

where $\alpha = \pi - \rho$.

17. The table below [quoted from Smith (1969, Section 21.17, Example 9)] gives the results of examining 100 schoolboys and 100 schoolgirls of similar ages for heart murmur:

	Murmur	No murmur	Total
Boys	58	42	100
Girls	46	54	100
Total	100	96	200

What are the approximate posterior odds that the proportion of boys with heart murmurs is at least 10 per cent greater than that of girls?

18. Suppose that $x \sim \mathrm{P}(10)$, i.e. x is Poisson of mean 10, and $y \sim \mathrm{P}(20)$. What is the approximate distribution of $2x + y$?

F.6 Further Exercises on Chapter 6

1. The sample correlation coefficient between length of wing and width of band on wing in females of ten populations of the butterfly *Heliconius charitonius* was determined for each of a number of populations. The results were as follows:

Population	1	2	3	4	5	6	7	8	9	10
Size	100	46	28	74	33	27	52	26	20	17
Correlation	0.29	0.70	0.58	0.56	0.55	0.67	0.65	0.61	0.64	0.56

Find an interval in which you are 90% sure that the correlation coefficient lies.

2. If you obtain a number of different sample correlation coefficients (with the same sample size in each case), all of which turn out to be positive, and combine them appropriately, is the resulting combined estimate less than or greater than their mean? Does the same answer hold without the restriction that they are all positive?

3. Suppoase you want a 95% HDR for the difference between two sample correlation coefficients each based on a sample of size n to be of total length 0.2. What sample size n do you need to take?

4. Show that the transformation $z = \tanh^{-1} r$ can also be expressed as

$$z = \tfrac{1}{2}\log\{(1+r)/(1-r)\}.$$

5. What prior for ζ does the Jeffreys' reference prior for ρ correspond to?

6. The following data consist of the yield (in hundredweights per acre) of wheat under 7 different levels of nitrogen (N):

N/acre	40	60	80	100	120	140	160
cwt/acre	15.9	18.8	21.6	25.2	28.7	30.4	30.7

Give an interval in which you are 90% sure that the yield of wheat will lie on a plot to which 90 units of nitrogen has been applied and give a similar interval in which the mean yield of all such plots lies.