Data provided: formula sheet

MAS223

The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2016–2017**

Statistical Modelling and Inference **2 hours 30 minutes**

*Candidates should attempt **ALL** questions.*
The maximum marks for the various parts of the questions are indicated.
The paper will be marked out of 90.

1 Let $\mathbf{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$ be a random vector with a bivariate normal distribution, with mean vector $\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.

(a) Write down the marginal distribution of X. ***(2 marks)***

(b) Let $U = X - Y$ and $V = X + 2Y + 3$. Find the mean vector and covariance matrix of the random vector $\mathbf{U} = \begin{pmatrix} U \\ V \end{pmatrix}$. ***(5 marks)***

2 Let (X, Y) be a bivariate random variable with probability density function

$$f_{X,Y}(x, y) = \begin{cases} k(x^2 + y) & \text{for } 0 < x < 1 \text{ and } 0 < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

where $k \in \mathbb{R}$ is a deterministic constant.

(a) Show that $k = \frac{6}{5}$. ***(2 marks)***

(b) Calculate $\mathbb{P}[X \leq Y]$. ***(3 marks)***

(c) Find the marginal probability density function $f_Y(y)$ of Y. ***(3 marks)***

3 Recall that the Beta function $B : (0, \infty) \times (0, \infty) \to \mathbb{R}$ is given by

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx.$$

(a) Show that $B(\alpha, 1) = \dfrac{1}{\alpha}$. ***(1 mark)***

(b) Let X be a random variable with the Beta distribution, $X \sim Be(8, 1)$, and let $Y = \sqrt[3]{X}$. Find the probability density function of Y and identify its distribution. ***(8 marks)***

4 Let X and Y be a pair of independent and identically distributed random variables. Let $U = X + Y$ and $V = X - Y$.

(a) Show that $\text{Cov}(U, V) = 0$. ***(3 marks)***

(b) Suppose, additionally, that $\lambda \in (0, \infty)$ and that X and Y both have the $Exp(\lambda)$ distribution.

Find the joint probability density function $f_{U,V}(u, v)$ of U and V, stating clearly the range of (u, v) for which it is non-zero. ***(8 marks)***

(c) With X, Y as in (b), are U and V independent? Justify your answer. ***(2 marks)***

5 Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a set of n independent identically distributed samples from the $Exp(\lambda)$ distribution, where $\lambda \in (0, \infty)$ is an unknown parameter and $n \geq 3$.

(a) Find the likelihood function $L(\lambda; \mathbf{x})$ and the log-likelihood function $\ell(\lambda; \mathbf{x})$. ***(4 marks)***

(b) Derive a formula for the maximum likelihood estimator $\hat{\lambda}$ of λ. ***(5 marks)***

(c) Sketch the log-likelihood function $l(\lambda; \mathbf{x})$, marking the location of $\hat{\lambda}$ clearly on your diagram. ***(2 marks)***

(d) Consider the set

$$R_2 = \Big\{\lambda \in (0, \infty) \;:\; |\ell(\lambda, \mathbf{x}) - \ell(\hat{\lambda}, \mathbf{x})| \leq 2\Big\}$$

Suggest why we might hope that values $\lambda \in R_2$ are good approximations to the true value of λ. ***(2 marks)***

If you wish, you may annotate your diagram from part (c) to help you answer part (d).

6 Consider the linear model

$$y_1 = \beta_0 + \beta_1 x_1 + \epsilon_1$$
$$y_2 = \beta_1 x_2 + \epsilon_2$$
$$y_3 = \beta_0 x_3 + \beta_1 + \epsilon_3$$

where the random errors $\epsilon_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$. The sample (x_i, y_i) is $(1, 0)$, $(0, 1)$, $(2, 2)$.

(a) Write down the design matrix of the above model. ***(2 marks)***

(b) Give the least squares estimates for β_0 and β_1. ***(4 marks)***

(c) Give an estimate for σ^2. ***(3 marks)***

(d) We wish to test $H_0 : \beta_0 = \beta_1$ versus $H_a : \beta_0 \neq \beta_1$. Perform the F-test and report the P-value in the form $P(F_{?,?} > ?)$. ***(6 marks)***

7 The dataset `florida` lists the total number of votes received by the candidates in the 2000 US presidential election, in each of the 67 counties in Florida. A simple linear regression model $y = \beta_0 + \beta_1 x + \epsilon$ was set up to test the relationship between votes cast for the conservative candidates Buchanan and Bush. Below is a partial `R` output.

```
> attach(florida)
> lm1<-lm(BUCHANAN ~ BUSH)
> summary(lm1)
```

`Coefficients:`

	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	45.29	54.48	0.831	0.409
BUSH	0.004917	0.0007644	6.432	1.73×10^{-8}

```
Residual standard error:  353.9 on 65 degrees of freedom
Multiple R-squared:  0.3889, Adjusted R-squared:  0.3795
F-statistic:  41.37 on 1 and 65 DF, p-value:  1.727 × 10^-8
```

(a) Use the information above to conduct the hypothesis test $H_0 : \beta_1 = 0.003$ versus $H_0 : \beta_1 \neq 0.003$. You must explain all the steps in your test. You might report the P-value in the form $P(t_? > ?)$ or $P(F_{?,?} > ?)$ depending on whether you use the t-test or the F-test. ***(4 marks)***

(b) Find a 95% confidence interval for β_0. You are given the following quantiles: ***(3 marks)***

$$t_{0.95,65} = 1.669, \ t_{0.975,65} = 1.997, \ t_{0.95,67} = 1.668,$$
$$t_{0.975,67} = 1.996, \ z_{0.95} = 1.65, \ z_{0.975} = 1.96$$

(c) Discuss the fit of the simple linear regression model. ***(2 marks)***

(d) Bush got 5413 votes in Bradford county. How many votes do we expect Buchanan to have got in that county? ***(1 mark)***

8 A study by Baumann and Jones of the Purdue University Education Department evaluates different teaching methods for reading comprehension in children. The methods are called "Basal", "Directed reading thinking activity (DRTA)" and "Strategies". Sixty-six children were randomly assigned to one of the three methods, and their reading comprehension was tested before and after instruction. The scores are stored as `prescore` and `postscore` respectively in the data set `reading`, along with the method of instruction stored as `group`. Below is a partial `R` output.

```
> attach(reading)
> lm0<-lm(postscore-prescore~1)
> lm1<-lm(postscore-prescore~group-1)
> anova(lm0,lm1)
```

```
Analysis of Variance Table
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(> F)
1	a	2841.4				
2	b	2457.7	c	d	e	0.01036

(a) Defining your notation carefully, write down the model considered in the above `R` output. ***(2 marks)***

(b) What is the test for which the F-value has been computed above? What is your conclusion for the test? ***(2 marks)***

(c) Compute `a, b, c, d, e`. ***(5 marks)***

9 24 patients (12 male and 12 female) were recruited in a study to test the relative effectiveness of three drugs A, B and C against malaria. It is believed that the sex of the patient is also a factor in the effectiveness of the drugs. Patients were randomly assigned to the three drugs in a *balanced design*. The drugs were administered immediately after detection of malaria and the recovery period (in days) was noted. Below is a table which lists the sample mean recovery period for each combination of sex and drug.

	A	B	C
Male	10.5	12	11
Female	11	11.5	12

(a) Defining your notation carefully, write down the most complex model it is sensible to consider for these data, and also state any necessary parameter constraints. ***(3 marks)***

(b) Sketch two plots which indicate the absence or presence of interaction between the factors. Discuss the plots. ***(3 marks)***

End of Question Paper

SOME DISCRETE DISTRIBUTIONS

Name	**Parameters**	**Genesis / Usage**	**Notation**	$p(x) = \mathbb{P}[X = x]$ **(and non-zero range)**	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	**Comments**
Uniform (discrete)	$k \in \mathbb{N}$	Set of k equally likely outcomes	$Unif(1, \ldots, k)$ (not standard)	$p(x) = 1/k$ $x = 1, ..., k$	$\frac{k+1}{2}$	$\frac{k^2-1}{12}$	Fair dice roll ($k = 6$)
Bernoulli trial	$\theta \in [0, 1]$	Experiment with two outcomes (typically, success = 1, fail = 0)	$Bernoulli(\theta)$	$p(x) = \theta^x(1-\theta)^{1-x}$ $x = 0, 1$	θ	$\theta(1-\theta)$	Coin toss
Binomial	$n \in \mathbb{N}$, $\theta \in [0, 1]$	Number of successes in n i.i.d. Bernoulli trials	$Bi(n, \theta)$	$p(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$ $x = 0, 1, 2, ..., n$	$n\theta$	$n\theta(1-\theta)$	Sampling with replacement $Bi(1, \theta) \equiv Bernoulli(\theta)$
Geometric	$\theta \in (0, 1]$	Number of failed i.i.d. Bernoulli trials before first success	$Geom(\theta)$	$p(x) = (1-\theta)^x\theta$ $x = 0, 1, 2, \ldots$	$\frac{1-\theta}{\theta}$	$\frac{1-\theta}{\theta^2}$	Alternative formulations might swap p and $1-p$, or use $X' = X + 1$ to include the successful trial
Negative Binomial (or Pascal)	$k \in \mathbb{N}$, $\theta \in (0, 1]$	Number of i.i.d. Bernoulli trials until k^{th} success	$NegBin(k, \theta)$ (not standard)	$p(x) = \binom{x-1}{k-1}\theta^k(1-\theta)^{x-k}$ $x = k, k+1, k+2, \ldots$	$\frac{k}{\theta}$	$\frac{k(1-\theta)}{\theta^2}$	Several alternative formulations exist.
Hypergeometric	$N \in \mathbb{N}$, $k \in \{0, \ldots, N\}$, $n \in \{0, \ldots, n\}$	Number of special objects in a random sample of n objects, from a population of N objects with k special objects	$HypGeom(N, k, n)$ (not standard)	$p(x) = \binom{n}{x}\binom{N-n}{k-x}/\binom{N}{n}$ $x = 0, ..., n$	$\frac{nk}{N}$	$n\frac{N-n}{N-1}\frac{k}{N}(1-\frac{k}{N})$	
Poisson	$\lambda \in (0, \infty)$	Counting events occurring 'at random' within space or time	$Poi(\lambda)$	$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $x = 0, 1, 2, ...$	λ	λ	

SOME CONTINUOUS DISTRIBUTIONS

Name	**Parameters**	**Genesis / Usage**	**Notation**	$f(x) =$ **p.d.f. (and non-zero range)**	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	**Comments**
Uniform (continuous)	$\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$	The uniform distribution for a continuous interval	$Unif(\alpha, \beta)$	$f(x) = \frac{1}{\beta - \alpha}$ $x \in (\alpha, \beta)$	$\frac{\alpha+\beta}{2}$	$\frac{(\beta-\alpha)^2}{12}$	Also written as $U[\alpha, \beta]$ and similarly for half-open intervals.
Normal	$\mu \in \mathbb{R}$, $\sigma \in (0, \infty)$	Empirically and theoretically (via CLT, etc.) a good model in many situations.	$N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $x \in \mathbb{R}$	μ	σ^2	$N(0,1) \equiv$ standard normal. $X \sim N(\mu, \sigma^2) \Rightarrow$ $aX + b \sim N(a\mu + b, a^2\sigma^2)$ Hence $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$
Exponential	$\lambda \in (0, \infty)$	Inter-arrival times of random events	$Exp(\lambda)$	$f(x) = \lambda e^{-\lambda x}$ $x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	Alternative parametrization: $\theta = \frac{1}{\lambda}$
Gamma	$\alpha, \beta \in (0, \infty)$	Lifetimes of ageing items, multi-inter-arrival times	$Ga(\alpha, \beta)$	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ $x > 0$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	Alternative parametrization: $\theta = 1/\beta$, $Ga(1, \lambda) \equiv Exp(\lambda)$, $Ga(n/2, 1/2) \equiv \chi^2_n$
Chi-squared	$n \in \mathbb{N}$	Squared (normally distributed) errors, statistical tests	χ^2_n	$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}$ $x > 0$	n	$2n$	$X^2_n \equiv Ga(n/2, 1/2)$ $X_i \sim N(0,1)$ i.i.d. $\Rightarrow \sum_{i=1}^n X_i^2 \sim \chi^2_n$
Beta	$\alpha, \beta \in (0, \infty)$	Quantities constrained to be within intervals	$Be(\alpha, \beta)$	$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ $x \in [0, 1]$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$Be(1,1) \equiv Unif(0,1)$
Cauchy	$a, b \in \mathbb{R}$	Heavy tailed	$Cauchy(a, b)$	$f(x) = \frac{1}{\pi b} \frac{b^2}{(x-a)^2+b^2}$ $x \in \mathbb{R}$	undefined	undefined	$Cauchy(0,1)$ is often called 'the' Cauchy distribution
Student t	$n \in \mathbb{N}$	Statistical tests	t_n	$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}$ $x \in \mathbb{R}$	0 if $n > 1$	$\frac{n}{n-2}$ if $n > 2$	$t_1 \equiv Cauchy(0,1)$ Can take $n \in (0, \infty)$
F	$\nu, \delta \in (0, \infty)$	Statistical tests	$F_{\nu,\delta}$	$f(x) = \frac{\nu^{\nu/2}\delta^{\delta/2}x^{\nu/2-1}}{B(\nu/2,\delta/2)(\nu x+\delta)^{(\nu+\delta)/2}}$ $x > 0$	$\frac{\delta}{\delta-2}$ if $\delta > 2$	$\frac{2\delta^2(\nu+\delta-2)}{\nu(\delta-2)^2(\delta-4)}$ if $\delta > 4$	If $X \sim \chi^2_\nu$ and $Y \sim \chi^2_\delta$ are independent then $\frac{X/\nu}{Y/\delta} \sim F_{\nu,\delta}$. If $T \sim t_\nu$ then $T^2 \sim F_{1,\nu}$. If $Z \sim Be(\alpha, \beta)$ then $\frac{\beta Z}{\alpha(1-Z)} \sim F_{2\alpha,2\beta}$.