Queen Mary
University of London

MTH5122: Statistical Methods

Duration: 2 hours

Date and time: 27 May 2016, 10:00–12:00

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

You should attempt ALL questions. Marks awarded are shown next to the questions.

Statistical functions provided by the calculator may be used provided that you state clearly where you have used them.

Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately. It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms it shall be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

Examiner(s): A. Gnedin

Question 1. [25 marks] Concentration of noradrenaline was measured in the brains of six rats that had been exposed to a toxic chemical (X-sample), and also in five unexposed control rats (Y-sample). The sample means and standard deviations are (measured in ng/g)

$$\bar{x} = 540.83,\ s_X = 66.11,\ \bar{y} = 444.20,\ s_Y = 69.64$$

(a) Suppose $\sigma_X^2 = \sigma_Y^2$. Test at a significance level of 5% the null-hypothesis $\mu_X = \mu_Y$ versus the alternative that the population mean is higher for the rats exposed to the chemical. State explicitly any other assumptions you make. **[10]**

(b) Show that at a significance level of 5% the equality of population variances σ_X^2 and σ_Y^2 cannot be rejected. **[10]**

(c) Suppose $\sigma_X^2 = \sigma_Y^2$. Find a 95% confidence interval for the difference of the population means $\mu_X - \mu_Y$. **[5]**

Table of $t_d(\alpha)$ (upper percentiles)

d, α	$8, 0.025$	$8, 0.05$	$9, 0.025$	$9, 0.05$	$10, 0.025$	$10, 0.05$
$t_d(\alpha)$	2.306	1.860	2.262	1.833	2.228	1.812

Table of $F_{d_2}^{d_1}(\alpha)$ (upper percentiles)

d_1, d_2, α	$5, 4, 0.025$	$4, 5, 0.025$	$5, 4, 0.05$	$4, 5, 0.05$	$6, 5, 0.05$	$5, 6, 0.05$	$6, 5, 0.025$
$F_{d_2}^{d_1}(\alpha)$	9.364	7.388	6.256	5.192	4.950	4.387	6.978

Question 2. [25 marks] Consider random variables X and Y with a marginal probability density function

$$f_X(x) = \begin{cases} 2\,e^{-2x}, & \text{for } 0 < x < \infty, \\ 0\,, & \text{otherwise,} \end{cases}$$

and a conditional probability density function

$$f_Y(y|X = x) = \begin{cases} e^{-(y-x)}, & \text{for } 0 < x < y < \infty, \\ 0\,, & \text{otherwise.} \end{cases}$$

(a) Name the distribution of X. **[3]**

(b) Determine the joint probability density function $f_{X,Y}$. **[3]**

(c) Sketch on the coordinate plane the support

$$I = \{(x, y) \in \mathbb{R}^2 : f_{X,Y}(x, y) > 0\}$$

of $f_{X,Y}$. **[3]**

(d) Find the marginal probability density function f_Y. **[5]**

(e) Calculate the conditional expectation $\mathbb{E}(Y|X = x)$ for $x > 0$. **[5]**

(f) Find the distribution of the random variable $X + Y$. You may use the method of generating functions, or the technique of density transforms, or any other appropriate method. **[6]**

Question 3. [25 marks] Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables such that $X_j \sim \chi^2_3$ (the chi-square distribution with three degrees of freedom), so $\mathbb{E}X_j = 3,\ \mathrm{Var}[X_j] = 6$, for $j = 1, 2, \ldots, n$.

(a) Name the distribution of $X_1 + X_2 + \cdots + X_n$ and specify its parameters. **[5]**

(b) Name the distribution of X_1/X_2 and specify its parameters. **[5]**

(c) Determine the moment generating function of $X_1 - X_2$. **[5]**

(d) Using Chebyshev's inequality show that $\mathbb{P}(|X_1 + X_2 + X_3 - 9| > 6) < 0.5$. **[5]**

(e) Apply the Central Limit Theorem to evaluate the limit value of the probability $\mathbb{P}(X_1 + X_2 + \cdots + X_n > 3n)$ as $n \to \infty$. **[5]**

Question 4. [15 marks] In a species of the plant snapdragon, individual plants can have red, pink or white flowers. By a Mendelian genetic model, self-pollinated pink-flowered plants should produce progeny that are red, pink and white in the ratio 1:2:1. A geneticist investigating the model observed 209 plants in the progeny with the following colours:

colour	red	pink	white
count	43	119	47

(a) Test the prediction of the Mendelian model at 5% significance level. **[5]**

(b) Give the definition of the P-value for a one-sided test. **[5]**

(c) Show that the P-value of the test is between 0.10 and 0.15 **[5]**

Table of $\chi^2_d(\alpha)$ (upper percentiles)

d, α	$2, 0.025$	$2, 0.975$	$2, 0.05$	$2, 0.10$	$2, 0.15$	$3, 0.05$	$4, 0.05$
$\chi^2_d(\alpha)$	7.38	0.05	5.99	4.61	3.79	7.81	9.49

Question 5. [10 marks] A popular belief is that sons are, on the average, taller than their fathers.

(a) Suggest a statistical test to verify the belief, given data on the heights of fathers and sons in a number of families with only one son. State clearly the assumptions necessary to apply the test. **[5]**

(b) Argue that the assumptions underlying the test are satisfied if the joint distribution of the height of son and the height of his father is a bivariate normal distribution. **[5]**

End of Paper.