Mathematical Tripos: Part IB SMP/Lent 2013

Statistics: Example Sheet 2 (of 3)

Comments and corrections to s.pitts@statslab.cam.ac.uk

1. Let X have density function $f(x;\theta) = \frac{\theta}{(x+\theta)^2}$, $x > 0$, where $\theta \in (0,\infty)$ is an unknown parameter. Find the likelihood ratio test of size 0.05 of $H_0 : \theta = 1$ against $H_1 : \theta = 2$, and show that the probability of Type II error is 19/21.

2. Let $X_1, X_2, \ldots, X_n$ be iid random variables, each with a Poisson distribution with parameter θ (and therefore with mean θ and variance θ). Find the form of the likelihood ratio test of $H_0 : \theta = 1$ against $H_1 : \theta = 1.21$. By using the Central Limit Theorem to approximate the distribution of $\sum_i X_i$, show that the smallest value of n required to make $\alpha = 0.05$ and $\beta \leq 0.1$ (where α and β are the Type I and Type II error probabilities) is somewhere near 212.

3. Let f_0 and f_1 be probability mass functions for $\mathbf{X} = (X_1 \ldots, X_n)$ on a countable set $\mathcal{X}^n$. State and prove a version of the Neyman–Pearson lemma for a size α test of $H_0 : f = f_0$ against $H_1 : f = f_1$, assuming that α is such that there exists a likelihood ratio test of exact size α.

4. Let $X \sim \text{Bin}(2,\theta)$ and consider testing $H_0 : \theta = \frac{1}{2}$ against $H_1 : \theta = \frac{3}{4}$. Find the possible values of α for which there exists a likelihood ratio test with size exactly α.

5. Let $X_1, \ldots, X_n$ be iid random variables each with a $N(\mu_0, \sigma^2)$ distribution, where μ_0 is known and σ^2 is unknown. Find the best (most powerful) test of size at most α for testing $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2$ for known σ_0^2 and σ_1^2 ($> \sigma_0^2$). Show that this test is a size α uniformly most powerful test for testing $H_0' : \sigma^2 \leq \sigma_0^2$ against $H_1' : \sigma^2 > \sigma_0^2$.

6. Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exponential}(\theta)$. Find the likelihood ratio test of size α of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ ($> \theta_0$) and derive an expression for the power function. Is the test uniformly most powerful for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$? Is it uniformly most powerful for testing $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$?

7. Let $X_1, \ldots X_n, Y_1, \ldots, Y_n$ be independent, with $X_1, \ldots, X_n \sim \text{Exponential}(\theta_1)$ and $Y_1, \ldots, Y_n \sim \text{Exponential}(\theta_2)$. Recalling the forms of the relevant MLEs from Sheet 1, show that the likelihood ratio of $H_0 : \theta_1 = \theta_2$ and $H_1 : \theta_1 \neq \theta_2$ is a monotone function of $|t - 1/2|$, where t is the observed value of the statistic T given by

$$T = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i + \sum_{i=1}^n Y_i}.$$

By writing down the distribution of T under H_0, express the likelihood ratio test of size α in terms of $|T - 1/2|$ and the percentage points of a beta distribution.
Hint: use Question 2 on Example Sheet 1.

8. A machine produces plastic articles (many of which are defective) in bunches of three articles at a time. Under the null hypothesis that each article has a constant (but unknown) probability θ of being defective, write down the probabilities $p_i(\theta)$ of a bunch having i defective articles, for $i = 0, 1, 2, 3$. In an trial run in which 512 bunches were produced, the numbers of bunches with i defective articles were 213 ($i = 0$), 228 ($i = 1$), 57 ($i = 2$) and 14 ($i = 3$). Carry out Pearson's chi-squared test at the 5% level of the null hypothesis, explaining carefully why the test statistic should be referred to the χ^2_2 distribution.

9. A random sample of 59 people from the planet Krypton yielded the results below.

		Eye-colour	
		1 (Blue)	2 (Brown)
Sex	1 (Male)	19	10
	2 (Female)	9	21

Carry out Pearson's chi-squared test at the 5% level of the null hypothesis that sex and eye-colour are independent factors on Krypton. Now carry out the corresponding test at the 5% level of the null hypothesis that each of the cell probabilities is equal to 1/4. Comment on your results.

10. Write down from lectures the model and hypotheses for a test of homogeneity in a two-way contingency table. By first deriving the MLEs under each hypothesis, show that the likelihood ratio and Pearson's chi-squared tests are identical to those for the independence test. Apply the homogeneity test to the data below from a clinical trial for a drug, obtained by randomly allocating 150 patients to three equal groups (so the row totals are fixed).

	Improved	No difference	Worse
Placebo	18	17	15
Half dose	20	10	20
Full dose	25	13	12

11. (a) Let $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$, and let A be an arbitrary $m \times n$ matrix. Prove directly from the definition that $A\mathbf{X}$ has an m-variate normal distribution. Show that $\mathrm{cov}(A\mathbf{X}) = A\Sigma A^T$, and that $A\mathbf{X} \sim N_m(A\boldsymbol{\mu}, A\Sigma A^T)$. Give an alternative proof that $A\mathbf{X} \sim N_m(A\boldsymbol{\mu}, A\Sigma A^T)$ using moment generating functions.

    (b) Let $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$, and let $\mathbf{X_1}$ denote the first n_1 components of $\mathbf{X}$. Let $\boldsymbol{\mu_1}$ denote the first n_1 components of $\boldsymbol{\mu}$, and let Σ_{11} denote the upper left $n_1 \times n_1$ block of Σ. Show that $\mathbf{X_1} \sim N_{n_1}(\boldsymbol{\mu_1}, \Sigma_{11})$.

$^+$12 In Question 3, does there exist a version of the Neyman–Pearson lemma when a likelihood ratio test of exact size α does not exist?