STATISTICS Part IB
Example Sheet 2 (of 3) Lent 2022

1. Let X have density function
$$f(x \mid \theta) = \frac{\theta}{(x+\theta)^2}, \quad x > 0,$$
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 $X \sim N(\mu, 1)$ where μ is unknown. Find the most powerful tests of sizes 0.05 and 0.01 for the following hypotheses:
(a) $H_0 : \mu = 0$ vs $H_1 : \mu = 4$.
(b) $H_0 : \mu = 4$ vs $H_1 : \mu = 0$.
Explain how to interpret your results when the realised value is $X(\omega) = 2.1$.

3. Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ be independent, with $X_1, \dots, X_n \sim \mathrm{Exp}(\theta_1)$ and $Y_1, \dots, Y_n \sim \mathrm{Exp}(\theta_2)$. 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
$$T = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i + \sum_{i=1}^n Y_i}.$$
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 quantiles of a beta distribution.

4. A machine produces biodegradable 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 a 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.

5. Let f_0 and f_1 be probability mass functions on a countable set $\mathcal{X}$. 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.

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

		Eye-colour	
		Blue	Brown
Sex	Male	19	10
	Female	9	21

Carry out a 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.

7. 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

8. Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Exp}(\theta)$. Find the likelihood ratio test of size α of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ where $\theta_1 > \theta_0$ 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 $H_1 : \theta > \theta_0$?

9. If $X \sim N(0,1)$ and $Y \sim \chi^2_n$ are independent, we say that $T = X/\sqrt{Y/n}$ has a t-distribution with n degrees of freedom and write $T \sim t_n$. Show that the probability density function of T is

$$f_T(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{(n\pi)^{1/2}} \frac{1}{(1 + t^2/n)^{(n+1)/2}}, \quad t \in \mathbb{R}.$$

10. Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, where σ^2 is unknown, and suppose we are interested in testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$. Letting $\overline{X} = n^{-1} \sum_{i=1}^n X_i$ and $S_{XX} = \sum_{i=1}^n (X_i - \overline{X})^2$, show that the likelihood ratio can be expressed as

$$L_X(H_0, H_1) = \left(1 + \frac{T^2}{n-1}\right)^{n/2},$$

where $T = \frac{n^{1/2}(\overline{X} - \mu_0)}{\{S_{XX}/(n-1)\}^{1/2}}$. Determine the distribution of T under H_0, and hence determine the size α likelihood ratio test.

11. Statisticians A and B obtain independent samples $X_1, \dots, X_{10}$ and $Y_1, \dots, Y_{17}$ respectively, both from a $N(\mu, \sigma^2)$ distribution with both μ and σ unknown. They estimate (μ, σ^2) by $(\overline{X}, S_{XX}/9)$ and $(\overline{Y}, S_{YY}/16)$ respectively, where for example, $\overline{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$ and $S_{XX} = \sum_{i=1}^{10} (X_i - \overline{X})^2$. Given that the values $\overline{X} = 5.5$ and $\overline{Y} = 5.8$, which statistician's estimate of σ^2 is more probable to have exceeded the true value by more than 50%? Find this probability (approximately) in each case.

12. **(Harder)** In Question 5, does there exist a version of the Neyman–Pearson lemma when a likelihood ratio test of exact size α does not exist?