MATH11400 Statistics 1 2010–11

Homepage http://www.stats.bris.ac.uk/%7Emapjg/Teach/Stats1/

Problem Sheet 9

Remember: when online, you can access the Statistics 1 data sets from an **R** console by typing
`load(url("http://www.stats.bris.ac.uk/%7Emapjg/Teach/Stats1/stats1.RData"))`

*1. For the data about fuel consumption on Problem Sheet 8, question 4, find a 90% confidence interval for the variance of fuel consumption per 100km for the population of cars of this type.

*2. Consider again the following failure-time data for the batch of 25 lamps (introduced on Problem Sheet 3), which you may assume is a simple random sample from an Exponential distribution with unknown parameter θ. The data is contained in the Statistics 1 data set `lamp`.

5.5	3.8	8.0	7.8	9.3	4.7	4.0	0.3	4.6	0.6	7.9	1.8	4.0
0.7	4.0	1.6	2.6	0.7	0.2	3.1	1.0	3.4	3.7	10.8	1.2	

(a) Use the method in §7.9 of your notes to find an equal-tailed 95% confidence interval for the unknown parameter θ based on this set of 25 observations.

(b) Let $X_1, \ldots, X_n$ be a simple random sample of size n from the Exp(θ) distribution. You may assume that $\mathrm{E}(1/\sum_{i=1}^n X_i) = \theta/(n-1)$ (for a derivation of this result, see the Solutions to question 4 from Problem Sheet 7). Use this result to find the average length of a 95% confidence interval for θ based on a random sample of size $n = 25$, expressing your answer as a multiple of the unknown parameter θ.

3. Consider again the opinion poll example, question 5 on Problem Sheet 6. Assume that a random sample of 1000 electors are interviewed and that 370 of those interviewed say that they support the govenment. Find a 99% confidence interval for the proportion of electors that support the govenment.

4. For the data about spatial-temporal reasoning of pre-school children on Sheet 8, question 5, under the assumption that the data are a simple random sample from a Normal distribution, construct a 95% confidence interval for the variance of the improvement in reasoning scores in the population.

*5. Assume the 25 observations below are a random sample from the Unif$(0, \theta)$ distribution.

1.41	0.11	0.61	4.06	2.81	4.23	2.68	4.43	2.98	4.15	0.10	4.04	5.57
2.04	4.44	5.48	1.53	0.10	4.82	5.99	2.35	0.07	3.24	5.83	1.57	

For the Unif$(0, \theta)$ distribution we saw earlier that the method of moments estimate $\hat{\theta}_{mom}$ and the maximum likelihood estimate $\hat{\theta}_{mle}$ were given by $\hat{\theta}_{mom} = 2\bar{X}$, where $\bar{X}$ is the sample mean, and $\hat{\theta}_{mle} = X_{(n)}$, where $X_{(n)} = \max(X_1, \ldots, X_n)$ is the sample maximum.

(a) Use the fact, that for a random sample of size n from the Unif$(0, \theta)$ distribution, $P(X_{(n)}/\theta \leq v) = v^n$ for $0 < v < 1$, to find values v_1 and v_2 such that $P(X_{(25)}/\theta < v_1) = 0.025$ and $P(X_{(25)}/\theta > v_2) = 0.025$. Hence, following the general idea seen in construction of other confidence intervals, but with different details, find an equal-tailed 95% confidence intervals for θ based on $\hat{\theta}_{mle}$.

(b) Find an equal-tailed 95% confidence intervals for θ based on $\hat{\theta}_{mom}$. [Hint: Use the Normal approximation to the distribution of $\bar{X}$ based on the Central Limit Theorem.] Comment on whether the interval you get is compatible with the data.

*6. A certain manufacturer produces packets of biscuits with a nominal weight of $200g$. You may assume that it is known from experience that the standard deviation of the weight of the packets is $4g$. To carry out a control check on the actual weight of the packets produced, an employee weighs 25 packets selected at random from a day's production and finds that the average weight of the sample is $\bar{x} = 202.275g$.

Let μ denote actual the mean weight of $200g$ packets produced by the manufacturer. Test the null hypothesis $H_0 : \mu = 200$ against the alternative $H_1 : \mu \neq 200$, using a test procedure with significance level $\alpha = 0.01$. For what range of significance levels would you reject H_0 in favour of H_1?

[Your answer should include a statement of any model assumptions, a brief description of your working at each stage of the test procedure including the p-value and the critical region for the test, and a summary of your conclusions.]

7. A random variable X is known to have a Normal distribution with mean μ and variance 25. To test the hypotheses

$$H_0 : \mu = 100 \quad \text{versus} \quad H_1 : \mu > 100$$

a test procedure is proposed which would take a simple random sample of size n from the population distribution of X and reject H_0 in favour of H_1 if the sample mean $\bar{x} > 102$, and otherwise accept H_0.

Find an expression in terms of the sample size n for the significance level α of this test procedure. Hence find the smallest sample size for which the significance level would be less than 0.05.