

## MTH6136 STATISTICAL THEORY : Exercise Sheet 1

Please try to attempt most of the questions **before** the exercise class. You should post your solution to **Questions 1&3 only** for marking in the yellow box on the ground floor of the Mathematics Building by **11 a.m. Thursday 21 January 2010**. Feedback on your work on the other question(s) will be given by me in my office hours, or in exercise classes.

1. Suppose that  $Y_1, Y_2, \dots, Y_{19}$  are independent  $N(\mu, \sigma^2)$  random variables and that the observed values of these random variables are  $y_1, y_2, \dots, y_{19}$ . The sample mean is  $\bar{y} = 23.6$  and the sample variance is  $s^2 = 15.4$ .
  - (a) Estimate  $\mu$  and  $\sigma^2$ , and hence calculate a 95% confidence interval for  $\mu$ .
  - (b) Test the null hypothesis that  $\mu = 20$  against the alternative that  $\mu \neq 20$ . Give a brief summary of your conclusion using non-mathematical language.
2. Consider the birth of twins. Let  $\phi$  denote the probability that they are identical, in which case they must be of the same gender. If they are fraternal, that is, not identical twins, each child is equally likely to be male or female independently of the other child.
  - (a) Justify the result that

$$P(\text{twins are of the same gender}) = \frac{1 + \phi}{2}.$$

- (b) Let  $Y$  be the number of pairs of twins of the same gender out of  $m$  pairs of twins. Then  $Y$  has a binomial distribution with parameters  $m$  and  $(1 + \phi)/2$ . Write down the likelihood for the single observation  $y$  of  $Y$ .
3. Write down the likelihood function for a random sample from a Poisson distribution with unknown mean  $\mu$ , and sketch it as a function of  $\mu$  (assume  $\bar{y} > 0$ ).

**MTH6136 STATISTICAL THEORY : Solutions to Exercise Sheet 1**

1. (a) The estimates of  $\mu$  and  $\sigma^2$  are  $\hat{\mu} = \bar{y} = 23.6$  and  $\hat{\sigma}^2 = s^2 = 15.4$ . Hence, a 95% confidence interval for  $\mu$  is given by

$$\bar{y} \pm t_{18,0.025} \times \frac{s}{\sqrt{n}} = 23.6 \pm 2.1009 \times \sqrt{\frac{15.4}{19}} = 23.6 \pm 1.89,$$

that is, [21.7,25.5].

- (b) The test statistic is

$$T = \frac{\bar{Y} - 20}{\sqrt{\frac{S^2}{n}}}.$$

Under  $H_0$ ,  $T \sim t_{n-1}$ . Now, the observed value is

$$t = \frac{23.6 - 20}{\sqrt{\frac{15.4}{19}}} = 4.00.$$

Comparing this with the  $t_{18}$  distribution, we see that we can reject  $H_0$  at the 0.1% level of significance, since  $t_{18,0.0005} = 3.9216$ . So, in conclusion, we have overwhelming evidence that the population mean is greater than 20.

2. (a) By the theorem of total probability, we may write

$$\begin{aligned} P(\text{same gender}) &= P(\text{same gender}|\text{identical})P(\text{identical}) \\ &\quad + P(\text{same gender}|\text{fraternal})P(\text{fraternal}) \\ &= 1 \times \phi + \frac{1}{2} \times (1 - \phi) \\ &= \frac{1 + \phi}{2}. \end{aligned}$$

- (b) By the definition of the likelihood, we have

$$\begin{aligned} L(\phi; y) = P(Y = y) &= \binom{m}{y} \left(\frac{1 + \phi}{2}\right)^y \left(1 - \frac{1 + \phi}{2}\right)^{m-y} \\ &= \binom{m}{y} \frac{(1 + \phi)^y (1 - \phi)^{m-y}}{2^m}. \end{aligned}$$

3. By the definition of the likelihood, we have

$$\begin{aligned} L(\mu; y_1, y_2, \dots, y_n) &= \prod_{i=1}^n P(Y_i = y_i) = \prod_{i=1}^n e^{-\mu} \mu^{y_i} / y_i! \\ &= e^{-n\mu} \mu^{\sum y_i} / \prod_{i=1}^n y_i! \\ &= a e^{-n\mu} \mu^{n\bar{y}}, \end{aligned}$$

where  $a$  a constant not depending on  $\mu$ , but on the data only. Hence  $L$  can be sketched easily as a function of  $\mu > 0$ , with the usual "positively skew distribution" shape.