THE ROYAL STATISTICAL SOCIETY

2010 EXAMINATIONS – SOLUTIONS

HIGHER CERTIFICATE

MODULE 3

BASIC STATISTICAL METHODS

Note. In accordance with the convention used in the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

(i) The null hypothesis is that there is no difference between the underlying population mean marks for the two groups. The alternative hypothesis is that there is a difference between the underlying population mean marks.

The pooled estimate of the assumed common variance is given by

$$s^2 = [(6 \times 312.90) + (10 \times 303.85)]/16 = 307.24.$$

The t statistic, with 16 df, is given by

$$\frac{61.64 - 40.71}{s\sqrt{\frac{1}{7} + \frac{1}{11}}} = 2.470\,.$$

For a two-sided test at the 5% significance level, the critical point from t_{16} is 2.120 and at the 1% level is 2.921.

So the observed value of 2.470 is significant at the 5% level but not at the 1% level. There is some evidence of a difference in the overall level of performance between the two groups – Group B appears to have performed better.

(ii) It is assumed that the both sets of marks may be thought of as random samples from Normal distributions with a common variance.

The two sample variances are very similar, which suggests that the assumption of a common underlying population variance is appropriate. The sample sizes are small, so the evidence is not very conclusive, but the data do not look as if they come from Normal distributions – both samples are negatively skewed and not obviously unimodal.

For the Wilcoxon rank sum test, it is assumed that both samples come from the same distribution apart from a possibly different location parameter, often identified with the median (i.e. they have the same shape of distribution, but possibly shifted between the two samples).

Solution continued on next page

(iii) The Wilcoxon rank sum test proceeds as follows.

The null hypothesis is that there is no difference between the underlying population location parameters/medians for the two groups. The alternative hypothesis is that there is a difference between the underlying population location parameters/medians

We rank all 18 data items:

Data	12	31	34	35	36	39	46	48	55	57	58	61	66	69	70	77	84	85
Ranks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	A	*A*	*A*	*A*	*B*	*B*	*B*	*B*	*A*	*A*	*B*	*A*	*B*	*B*	*B*	*B*	*B*	*B*

The rank sum for the smaller sample, which is the sample for *A*, is

$$1 + 2 + 3 + 4 + 9 + 10 + 12 = 41$$

We again use a 5% significance level.

The required test is two-sided. Using a 5% significance level, the critical point for $n_1 = 7$, $n_2 = 11$ is 44 (note that the table headed "0.025 level" in the Society's *Statistical tables for use in examinations* must be used).

Thus the calculated sample statistic of 41 is significant at the 5% level. There is some evidence of a difference in the overall level of performance between the two groups – Group B appears to have performed better.

<u>Higher Certificate, Module 3, 2010. Question 2</u>

(i) The binomial distribution with parameters 10 and θ, i.e. the $B(10, \theta)$ distribution.

(ii) The null hypothesis is $\theta = ½$. The alternative hypothesis is $\theta \neq ½$.

(iii) Using the table of the binomial cumulative distribution function (cdf) with $n = 10$, $\theta = 0.50$, $x = 2$, we find that under the null hypothesis, i.e. if in fact $X \sim B(10, ½)$, $P(X \leq 2) = 0.0547$.

Since we are dealing with a two-sided alternative hypothesis, the p-value corresponding to the observed value $x = 2$ is given by $2 \times 0.0547 = 0.1094$. Since this does not correspond with significance at the 5% level (or even at the 10% level), we conclude that there is no real evidence that the proportion of Type A manuscripts originally produced differed from the proportion of Type B manuscripts produced.

(iv) Using the Normal approximation to the binomial distribution, the approximate distribution is Normal with mean 100θ and variance $100\theta(1 - \theta)$.

(v) Under the null hypothesis $\theta = ½$, the number X of Type A manuscripts in the sample of size 100 is given approximately by $X \sim N(50, 25)$.

The standardised Normal variate for calculating the value of the cdf corresponding to the observed value $x = 39$ (and using a continuity correction) is therefore given by $z = (39.5 - 50)/\sqrt{25} = -2.1$.

Using the table of the Normal cdf with $z = 2.1$, we find $\Phi(2.1) = 0.9821$. Therefore, corresponding to our two-sided alternative hypothesis, we have a two-tailed test with p-value $= 2 \times (1 - 0.9821) = 0.0358$. So the result is significant at the 5% level, to the effect that the proportions differed – it appears that more manuscripts of Type B were produced than of Type A.

Higher Certificate, Module 3, 2010. Question 3

(i) If p_i denotes the probability that a random digit takes the value i, then

$$p_i = 1/10 \quad (\text{for } i = 0, 1, 2, \ldots, 9).$$

Hence the expected frequency for i is given by

$$E_i = 100p_i = 10 \quad (\text{for } i = 0, 1, 2, \ldots, 9).$$

(ii) The appropriate test statistic is a chi-squared statistic. It is calculated using the formula

$$X^2 = \Sigma(O_i - E_i)^2/E_i$$

giving

$$X^2 = [0^2 + 3^2 + 2^2 + 2^2 + 0^2 + 3^2 + 3^2 + 6^2 + 5^2 + 4^2]/10 = 11.2.$$

This is referred to the χ^2 distribution with 9 degrees of freedom.

The 5% point of this distribution is 16.919 and the 10% point is 14.684. Hence there is no real evidence judged at the 5% level (or even at the 10% level) to reject the assumed discrete uniform distribution for the digits. So there are no grounds here to call into question the adequacy of the pseudorandom number generator.

(iii) From the table of the Normal cumulative distribution function, the required probability p is given by $p = 2 \times (1 - 0.9772) = 0.0456$.

(iv) The observed proportion is $\hat{p} = 54/1000 = 0.054$.

Thus the test statistic for testing the null hypothesis that $p = 0.0456$ against the two-sided alternative that $p \neq 0.0456$ is

$$\frac{0.054 - 0.0456}{\sqrt{\dfrac{0.0456(1-0.0456)}{1000}}} = 1.273$$

We refer this to the N(0, 1) distribution. The critical point for a two-tailed test at the 5% significance level is 1.96 and at the 10% level is 1.645. So there is no real evidence for rejecting the null hypothesis and therefore no grounds here to suggest that the pseudorandom number generator is not performing as it should.

Higher Certificate, Module 3, 2010. Question 4

Part (a)

If independent random samples of a given size n from a given distribution are taken repeatedly, then the sampling distribution of a sample statistic is the probability distribution of how it varies from sample to sample.

In the given case:

the sampling distribution of $\overline{X}$ is Normal with mean μ and variance σ^2/n;

the sampling distribution of S^2 is given by $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.

Part (b)

(i) A 95% confidence interval for μ is given by

$\bar{x} \pm t_{99}(0.025)\, s/\sqrt{100}$

where $t_{99}(0.025)$ denotes the double-tailed 5% point of the t_{99} distribution which is here taken as 1.984 (the corresponding tabulated point for the t_{100} distribution) [**Note.** For a sample of this size, it would also be reasonable to use the N(0, 1) distribution with critical point 1.96.]

i.e. by $453.08 \pm (1.984)(5.42)/10$

i.e. by 453.08 ± 1.075

i.e. it is (452.00,454.16).

(ii) A 95% confidence interval for the population variance of the weights is based on the χ^2_{99} distribution, using the result $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.

The lower and upper 2½% points of this distribution are 73.36 and 128.42. [**Note.** Any reasonable interpolation in the table between 90 and 100 df was allowed in the examination.] . Hence the interval is given by

$$\left(\frac{99s^2}{128.42}, \frac{99s^2}{73.36}\right) = \left(\frac{2908.26}{128.42}, \frac{2908.26}{73.36}\right) = (22.65,\ 39.64).$$

Solution continued on next page

(iii) Let s_1^2 denote the sample variance for Brand B and s_2^2 the sample variance for Brand A. Under the null hypothesis that the population variances for Brand B and Brand A are the same, the distribution of s_1^2/s_2^2 is $F_{19,99}$.

We carry out a one-tailed test to examine the one-sided alternative hypothesis that the population variance for Brand B is greater than that for Brand A.

The test statistic s_1^2/s_2 has the value $11.15^2/5.42^2 = 4.23$.

The upper 0.1% point of the $F_{19,99}$ distribution from interpolation in the table is given (approximately) by 2.64 (or 2.63). [**Note.** Any reasonable interpolation in the table between the values for 18 and 24 and for 90 and 100 was allowed in the examination. Alternatively, it can be argued that the observed value is clearly in excess of all 0.1% critical points in this region of the table.]

The test statistic is well above this value. There is overwhelming evidence that the underlying variance is greater for Brand B than for Brand A.