PAPER A

Advanced Higher Statistics Practice Prelim

1. A Gardening Club went in three buses to a Garden Festival. The frequency distribution of the ages of those on the trip was as follows:

age	60	61	62	63	64	65	66	67	68	69	70	75	77
frequency	10	12	11	8	16	14	9	13	18	15	6	1	1

Would either of the two eldest people be considered as an outlier for this distribution? **3**

2. A bag contains 4 blue and 2 red counters. 2 counters are drawn at random without replacement. The random variable X is the number of blue counters drawn.

a) Find the probability distribution for X. **2**

b) Find $E(X)$ and $V(X)$. **2**

c) Evaluate the mean and standard deviation of $12 - 2X$ **3**

3. The probability that it will rain on any given day is 0.45. Calculate the probability that, in a two week period, it will rain

a) On more than 10 days **2**

b) Between 4 and 8 days (inclusive) **1**

4. 2.5% of the population carry a gene which can cause a degenerative disease. There is a test for the gene which correctly detects it in a carrier with probability 0.96 but which incorrectly detects it in a non-carrier with probability 0.02.

a) Calculate the probability that a randomly selected person who has been given the test will test positive for the gene. **2**

b) Calculate the probability that a person, found by the test to be carrying the gene, does in fact carry it. **3**

5. In the manufacture of a rope, it is found that there is a mean of 1.5 flaws per metre in the fabric of the rope and a mean of 3 flaws per metre in the weaving process. Both types of flaw follow independent Poisson distributions. Find the probability that there are

a) Exactly 6 fabric flaws in 5 metres of rope. **2**

b) More than 8 weaving flaws in 2 metres of rope. **2**

c) Using an approximation, find the probability that there are less than 50 total flaws in 10 metres of rope. **5**

PAPER A

6. The weights of a batch of tomatoes are normally distributed with mean 120 g and variance 300g. They are graded according to their weight.

a) Grade A tomatoes must have a weight of at least 135g. What percentage of the tomatoes will be grade A? **3**

b) 8% of the tomatoes are Grade D (the lowest grade). What is the maximum weight of a grade D tomato? **4**

7. Police records show that in 2014, 31% of drink driving offences were by males aged under 25. In 2015, a random sample of 10 offences from each of the 5 police forces in the country was chosen and 22 of these offences were by males aged under 25.

a) State how bias might have arisen in this sampling method. **1**

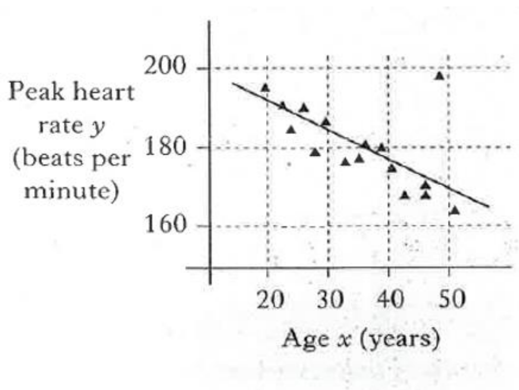
b) Calculate a 90% confidence interval for the population proportion of offences by males aged under 25 in 2015. **3**

c) Is there evidence to suggest the proportion of offences by males aged under 25 has changed from 2014 to 2015? **2**

8. It is known that in a sack of mixed grass seeds, 35% are ryegrass. Use a suitable approximation to calculate the probability that in a sample of 400 seeds that there are between 120 and 150 ryegrass seeds (inclusive). **6**

9. As part of a study on intensive exercise, a sports scientist recorded the peak heart rates of a random selection of sixteen volunteers of different ages who took regular exercise. The linear regression equation was calculated for the data shown in the scatter diagram and found to be

$$y = 209 - 0\cdot727x.$$



However, after considering the scatter diagram for the data, it was realised that one piece of data had been misrecorded and this volunteer's data was ignored.

a) State the approximate age of the volunteer whose data was ignored. **1**

b) Calculate the new regression equation using the values

$\sum x = 509$, $\sum x^2 = 18477$, $\sum y = 2738$, $\sum y^2 = 501192$, $\sum xy = 91694$ **2**

c) Comment on the difference this makes to the prediction for the average peak heart rate of a 45 year old volunteer. **2**

PAPER A

10. A machine fills cartons of orange juice with volume normally distributed with mean 1015 ml and standard deviation 4 ml. It is suspected that the machine may be malfunctioning and that the mean volume has changed. A sample of 10 cartons is taken and the sample mean is calculated to be 1012.2 ml.

    a) Test, at the 5% significance level, whether the mean volume has changed, stating any assumption made. **6**

    b) Use the sample mean to find a 95% confidence interval for the mean volume and explain how this interval confirms the answer to a). **4**

11. If the continuous random variable X ~ U(12, 36), calculate $P(20 \leq X \leq 26)$. **2**

12. In a newspaper article entitled "Sweet-toothed men are Britain's true chocoholics", it was reported that "more than half (53%) of adults regularly snack on chocolate". The survey on which the article was based involved a sample of 1000 men and 531 members of the sample indicated that they snacked regularly on chocolate.

    a) State two potential problems in attempting to obtain accurate information on snacking habits from a representative sample of 1000 British men. **2**

    b) State an advantage and disadvantage of carrying out the survey by phone. **2**

13. Each member of the Scottish 4x400m relay team can run their leg in a time normally distributed with mean 46.3 seconds and standard deviation 0.6 seconds.

    a) Find the mean and standard deviation for the total time taken by the team for the whole race. **2**

    b) The English team can run the whole race in a time normally distributed with mean 3 minutes 4.5 seconds and standard deviation 1.3 seconds. Find the probability that the Scottish team will beat the English team in a race. **5**

14. A machine produces a component for a car, the length of which is normally distributed with mean 136.5 cm.

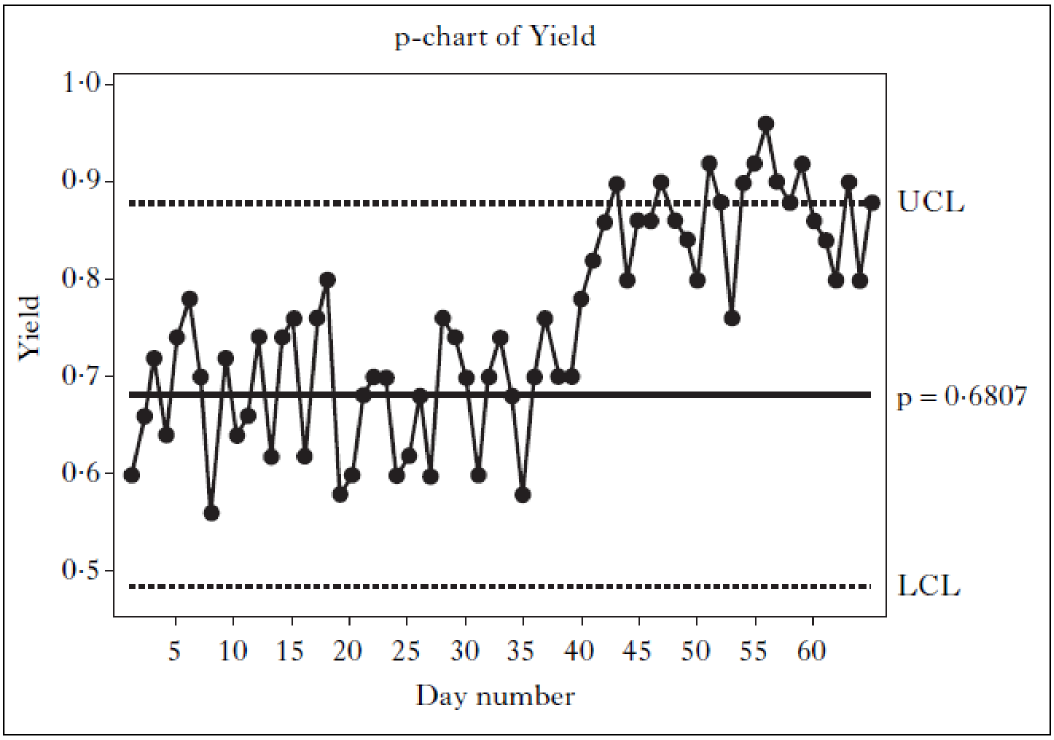
    a) If the standard deviation of the lengths is 0.25 cm, find the interval, symmetric about the mean, within which 98% of the lengths lie. **4**

    b) The car manufacturer reduces tolerances, demanding that 99% of the lengths must lie between 136.2 cm and 136.8 cm. To achieve this a new machine is required, producing lengths with a smaller standard deviation. What should this standard deviation be? **4**

PAPER A

15.

A10. In a manufacturing operation, daily samples of items were checked for nonconformities, the sample size being constant. The proportion of items free from nonconformities is referred to as Yield. The p-chart of Yield for 60 days production is shown below with 3-sigma limits. Process changes were made after 40 days.



(*a*) Given that the sample size is 50, calculate the upper and lower control limits.

(*b*) Explain how the chart provides evidence that the process changes have had a beneficial effect on Yield.

It was decided to chart the data from day 41 onwards in a new p-chart with limits based on the data for days 41 to 60 inclusive. The mean Yield for days 41 to 60 inclusive was 0·8648.

(*c*) Show that this would result in an impossible upper limit.

(*d*) By solving an inequality, determine the smallest sample size that would lead to a viable upper limit for monitoring the modified process.

3, 1, 2, 3