# OCR
**RECOGNISING ACHIEVEMENT**

## ADVANCED GCE UNIT

**4734/01**

## MATHEMATICS

Probability & Statistics 3

**FRIDAY 12 JANUARY 2007**

Morning

Time: 1 hour 30 minutes

Additional Materials: Answer Booklet (8 pages)
List of Formulae (MF1)

---

## INSTRUCTIONS TO CANDIDATES

*   Write your name, centre number and candidate number in the spaces provided on the answer booklet.
*   Answer **all** the questions.
*   Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
*   You are permitted to use a graphical calculator in this paper.

## INFORMATION FOR CANDIDATES

*   The number of marks is given in brackets [ ] at the end of each question or part question.
*   The total number of marks for this paper is 72.

## ADVICE TO CANDIDATES

*   Read each question carefully and make sure you know what you have to do before starting your answer.
*   **You are reminded of the need for clear presentation in your answers.**

---

This document consists of **4** printed pages.

OCR is an exempt Charity  **[Turn over**

1    The marks obtained by a randomly chosen student in the two papers of an examination are denoted by the random variables $X$ and $Y$, where $X \sim N(45, 81)$ and $Y \sim N(33, 63)$. The student's overall mark for the examination, $T$, is given by $T = X + \lambda Y$, where the constant $\lambda$ is chosen such that $E(T) = 100$.

   (i)   Show that $\lambda = \frac{5}{3}$.                                                                   [2]

   (ii)  Assuming that $X$ and $Y$ are independent, state the distribution of $T$, giving the values of its parameters.                                                                                    [3]

   (iii) Comment on the assumption of independence.                                                  [1]


2    The continuous random variable $X$ takes values in the interval $0 \leqslant x \leqslant 3$ only with probability density function f. The graph of $y = f(x)$ consists of the two line segments shown in the diagram.
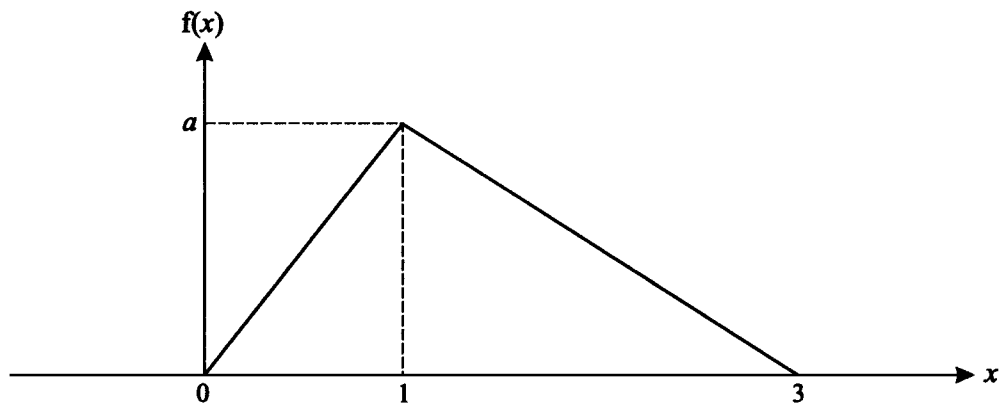


   (i)   Show that $a = \frac{2}{3}$.                                                                          [1]

   (ii)  Find the equations of the two line segments.                                                    [3]

   (iii) Hence write down the probability density function of $X$.                                      [1]

   (iv)  Find $E(X)$.                                                                                          [4]


3    A new treatment of cotton thread, designed to increase the breaking strength, was tested on a random sample of 6 pieces of a standard length. The breaking strengths, in grams, were as follows.

   17.3    18.4    18.6    17.2    17.5    19.3

   The breaking strengths of a random sample of 5 similar pieces of the thread which had not been treated were as follows.

   18.6    17.2    16.3    17.4    16.8

   A test of whether the treatment has been successful is to be carried out.

   (i)   State what distributional assumptions are needed.                                              [2]

   (ii)  Carry out the test at the 10% significance level.                                                [8]

**4** A machine is set to produce metal discs with mean diameter 15.4 mm. In order to test the correctness of the setting, a random sample of 12 discs was selected and the diameters, $x$ mm, were measured. The results are summarised by $\Sigma x = 177.6$ and $\Sigma x^2 = 2640.40$. Diameters may be assumed to be normally distributed with mean $\mu$ mm.

(i) Find a 95% confidence interval for $\mu$. [6]

(ii) Test, at the 5% significance level, the null hypothesis $\mu = 15.4$ against the alternative hypothesis $\mu < 15.4$. [4]

**5** Each person in a random sample of 1200 people was asked whether he or she approved of certain proposals to reduce atmospheric pollution. It was found that 978 people approved. The proportion of people in the whole population who would approve is denoted by $p$.

(i) Write down an estimate $\hat{p}$ of $p$. [1]

(ii) Find a 90% confidence interval for $p$. [4]

(iii) Explain, in the context of the question, the meaning of a 90% confidence interval. [2]

(iv) Estimate the sample size that would give a value for $\hat{p}$ that differs from the value of $p$ by less than 0.01 with probability 90%. [4]

**6** The lifetime of a particular machine, in months, can be modelled by the random variable $T$ with probability density function given by

$$f(t) = \begin{cases} \dfrac{3}{t^4} & t \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Obtain the (cumulative) distribution function of $T$. [2]

(ii) Show that the probability density function of the random variable $Y$, where $Y = T^3$, is given by $g(y) = \dfrac{1}{y^2}$, for $y \geq 1$. [6]

(iii) Find $E(\sqrt{Y})$. [3]

**[Question 7 is printed overleaf.]**

**7** It is thought that a person's eye colour is related to the reaction of the person's skin to ultra-violet light. As part of a study, a random sample of 140 people were treated with a standard dose of ultra-violet light. The degree of reaction was classified as None, Mild or Strong. The results are given in Table 1. The corresponding expected frequencies for a $\chi^2$ test of association between eye colour and reaction are shown in Table 2.

**Table 1**
**Observed frequencies**

| | | Eye colour | | | |
| | | Blue | Brown | Other | Total |
|---|---|---|---|---|---|
| | None | 12 | 17 | 10 | 39 |
| Reaction | Mild | 31 | 21 | 11 | 63 |
| | Strong | 22 | 4 | 12 | 38 |
| | Total | 65 | 42 | 33 | 140 |

**Table 2**
**Expected frequencies**

| | | Eye colour | | |
| | | Blue | Brown | Other |
|---|---|---|---|---|
| | None | 18.11 | 11.70 | 9.19 |
| Reaction | Mild | 29.25 | 18.90 | 14.85 |
| | Strong | 17.64 | 11.40 | 8.96 |

**(i) (a)** State suitable hypotheses for the test. [2]

**(b)** Show how the expected frequency of 18.11 in Table 2 is obtained. [1]

**(c)** Show that the three cells in the top row together contribute 4.53 to the calculated value of $\chi^2$, correct to 2 decimal places. [3]

**(d)** You are given that the total calculated value of $\chi^2$ is 12.78, correct to 2 decimal places. Give the smallest value of $\alpha$ obtained from the tables for which the null hypothesis would be rejected at the $\alpha\%$ significance level. [2]

**(ii)** Test, at the 5% significance level, whether the proportions of people in the whole population with blue eyes, brown eyes and other colours are in the ratios 2 : 2 : 1. [7]

---

**1**    (i) $E(T) = E(X) + \lambda E(Y)$      M1        Use $E(X + \lambda Y)$

         $\Rightarrow 100 = 45 + 33\lambda$

         $\Rightarrow \lambda = {}^5/_3$ AG              A1    **2**     aef

-------------------------------------------------------------------------------------------------

     (ii)     $Var(T) = Var(X) + ({}^5/_3)^2 Var(Y)$     M1

              $= 256$                   A1

              $T \sim N(100, 256)$      B1√   **3**     ft variance

-------------------------------------------------------------------------------------------------

     (iii)    Same student for $X$ and $Y$ so

            independence unlikely.      B1    **1**     Sensible reason

---

**2**    (i)    Use $3a/2 = 1$             B1    **1**     Or similar

-------------------------------------------------------------------------------------------------

     (ii)    $y = {}^2/_3 x$                B1

          $y = 1 - {}^1/_3 x$        M1A1  **3**    M1 for correct gradient

                                           B1M1A0 if not $y=...$

-------------------------------------------------------------------------------------------------

     (iii)    $f(x) = \begin{cases} \dfrac{2}{3}x & 0 \leq x \leq 1 \\[2mm] 1 - \dfrac{1}{3}x & 1 < x \leq 3. \end{cases}$

                                             B1√    **1**     ft (ii)

-------------------------------------------------------------------------------------------------

     (iv)    $\displaystyle\int_0^1 \frac{2}{3}x^2 dx + \int_1^3 (x - \frac{1}{3}x^2) dx$      M1        One correct, with limits

          $\left[\dfrac{2}{9}x^3\right]_0^1 + \left[\dfrac{1}{2}x^2 - \dfrac{1}{9}x^3\right]_1^3$     A1√A1√     ft from similar f

          $= 4/3$                              A1    **4**     aef

---

**3**    (i) Assumes breaking strengths have normal

      normal distributions                     B1

      Equal variances                       B1    **2**

-------------------------------------------------------------------------------------------------

     (ii) $H_0 : \mu_T = \mu_U$, $H_1 : \mu_T > \mu_U$        B1        For both hypotheses

          where $\mu_T, \mu_U$ are means for

          treated and untreated thread.

     $\bar{x}_T = 18.05, \quad \bar{x}_U = 17.26$       B1        May be implied below by 0.79

     $s_T^2 = 0.715, \quad s_U^2 = 0.738$       B1        Allow biased, 0.596, 0.590 if $s^2$

     $s^2 = (5 \times 0.715 + 4 \times 0.738)/9$     M1        $= (6 \times 0.596 + 5 \times 0.590)/9$

     EITHER:$(18.05 - 17.26)/[s\sqrt{1/5 + 1/6}]$   M1      With pooled variance est.

       $= 1.532$                                 A1

     Compare correctly with 1.383        M1

     Reject $H_0$ and accept there is sufficient

     evidence that mean has increased so that

     the treatment has been successful.      A1√        Conclusion in context. Ft 1.532

     OR: $\bar{X}_T - \bar{X}_U \geq ks\sqrt{1/5 + 1/6} ; = 0.713$   M1A1      Allow > or =

     $0.79 > 0.713$, reject $H_0$ etc        M1A1√ **8**    Or equivalent. Ft 0.713

---

| | | | | |
|---|---|---|---|---|
| 4 | (i) $s^2 = \frac{1}{11}(2604.4 - 177.6^2/12)$ | M1 | | aef |
| | $= 1.0836\ldots$ | A1 | | |
| | Use $\bar{x} \pm t\sqrt{\dfrac{s^2}{12}}$ | M1 | | |
| | $t = 2.201$ | B1 | | |
| | $\bar{x} = 177.6/12 = 14.8$ | A1 | | |
| | $(14.14, 15.46)$, $(14.1, 15.5)$ | A1 | **6** | |

----------------------------------------------------------------------------------------------------

| | | | | |
|---|---|---|---|---|
| | (ii) EITHER:$(14.8 - 15.4)/(\sqrt{(s^2/12)})$ | M1 | | With their variance |
| | $= -1.997$ | A1 | | |
| | Compare correctly with $-1.796$ | M1 | | |
| | Reject $H_0$ and accept that there is | | | |
| | evidence that the mean is less than 15.4 | A1√ | | **I**n context. Ft $-1.997$ |
| | OR: $\bar{X} - 15.4 \le -k\sqrt{\dfrac{s^2}{12}}$ ; $\bar{X} \le 14.86$ | M1A1 | | Allow $<$ or $=$ |
| | $14.8 < 14.86$, reject $H_0$ etc | M1A1√ | **4** | Or equivalent. Ft 14.86 |

_____

| | | | | |
|---|---|---|---|---|
| 5 | (i) $978/1200 = 0.815$ | B1 | **1** | |

----------------------------------------------------------------------------------------------------

| | | | | |
|---|---|---|---|---|
| | (ii) Use $\hat{p} \pm z\sqrt{\dfrac{\hat{p}(1-\hat{p})}{1200}}$ | M1 | | Reasonable variance |
| | $z = 1.645$ | B1 | | |
| | $\sqrt{(0.815\times0.185/1200)}$ | A1√ | | ft $\hat{p}$     Allow 1199 |
| | $(0.797, 0.833)$ | A1 | **4** | Interval |

----------------------------------------------------------------------------------------------------

| | | | | |
|---|---|---|---|---|
| | (iii) If a large number of such samples were | | | |
| | taken, $p$ would be contained in about 90% | | | B1 if idea correct but badly |
| | of the confidence intervals. | B2 | **2** | expressed. |

----------------------------------------------------------------------------------------------------

| | | | | |
|---|---|---|---|---|
| | (iv) $1.645\sqrt{(0.815\times0.185/n)} = 0.01$ | M1 | | Allow one error; $>$ or $<$ |
| | | A1 | | All correct |
| | $n = 1.645^2(0.815\times0.185)/0.01^2$ | M1 | | Correct procedure for sim equ |
| | $= 4080$ | A1 | **4** | Integer rounding to 4100 |

_____

6     (i) $\displaystyle\int_1^t \frac{3}{x^4}\,dx$            M1            Any variable

$$F(t) = \begin{cases} 1 - \dfrac{1}{t^3} & t \geq 1, \\ (0 & \text{otherwise.)} \end{cases}$$    A1     **2**

-----------------------------------------------------------------------------------------------------------------------------

    (ii) G($y$) = P( $Y \leq y$)           M1
           = P($T \leq y^{1/3}$)        A1
           = F($y^{1/3}$)          M1
           = 1 – 1/$y$           A1 √       ft F($t$)
       g($y$) = G'($y$)           M1
           = 1/$y^2$, $y \geq 1$    AG      A1     **6**

-----------------------------------------------------------------------------------------------------------------------------

    (iii) EITHER $\displaystyle\int_1^\infty \frac{\sqrt{y}}{y^2}\,dy$       OR $\displaystyle\int_1^\infty \frac{3t^{\frac{3}{2}}}{t^4}\,dt$    M1

      $\left[ -2y^{-1/2} \right]_1^\infty$        $\left[ -2t^{-3/2} \right]_1^\infty$       B1

          = 2                            A1     **3**

---

7     (i)(a)   $H_0$: Eye colour and reaction are not
         associated.                 B1         Or equivalent (independent, or
                                                  unrelated)
         $H_1$: Eye colour and reaction are
         associated                  B1    **2**
     (b)    65×39/140              B1    **1**

     (c)    $6.11^2/18.11 + 5.3^2/11.7 + 0.81^2/9.19$    M1       Or equivalent ; one correct
         2.061 + 2.401+ 0.071           A1         At least 3 dp here
         4.533, 4.53 AG          A1    **3**    But accept from 2 dp

     (d)   $v = 4$                        B1         Stated or implied
         Use tables to obtain $\alpha = 2\frac{1}{2}$      B1    **2**

---------------------------------------------------------------------------------------------------------------

    **(ii)** $H_0$: $p_{BL} = p_{BR} = 0.4$, $p_O = 0.2$      B1        Or in words, in terms of
     ($H_1$: At least two prob. not as above)            probs or proportions
    E values 56      56       28                M1A1
    $\chi^2 = 9^2/56 + 14^2/56 + $     $8^2/28$       M1
      = 5.839                       A1         Accept 5.84
    Compare correctly with 5.991          M1         M1A0 if 5.991 seen and
    Accept that sample is consistent with                        consistent conclusion but       the
hypothesis.                        A1√    **7**    no explicit comparison
    SR: If three tests for $p$ then count only
    $p_{BR} = 0.4$.
    (42/140 – 0.4)/$\sigma$                M1
    $\sigma = \sqrt{(0.4 \times 0.6/140)}$; -2.415       A1A1
    Compare with −1..96; conclusion in context    M1A1        Max 6/7 (with $H_0$)