**Paper 2 Option 2E**

**Further Statistics 1 Mark Scheme (Section A)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Question** | **Scheme** | **Marks** | **AOs** |
| **1(a)** | H0: There is no association between language and gender | B1 | 1.2 |
|  | **(1)** |  |
| **(b)** | \* | B1\*cso | 1.1b |
|  | **(1)** |  |
| **(c)** | |  |  |  |  |  | | --- | --- | --- | --- | --- | |  |  | Language | | | | Expected frequencies | | French | Spanish | Mandarin | | Gender | Male | 26.43… | 23.4 | 15.16… | | Female | 34.56… | [30.6] | 19.83… |     Awrt **3.6/3.7** | M1  M1  A1 | 2.1  1.1b  1.1b |
|  | **(3)** |  |
| **(d)** | Degrees of freedom (3 – 1)(2 – 1) → Critical value | M1 | 3.1b |
| As < 9.210, the null hypothesis is not rejected | A1 | 2.2b |
|  | **(2)** |  |
| **(e)** | Still not rejected since | B1 | 2.4 |
|  | **(1)** |  |
| **(8 marks)** | | | |
| **Notes:** | | | |
| **(a)**  **B1:** For correct hypothesis in context | | | |
| **(b)**  **B1\*:** For a correct calculation leading to the given answer and no errors seen | | | |
| **(c)**  **M1:** For attempt at  to find expected frequencies  **M1:** For applying  **A1:** awrt 3.6 or 3.7 | | | |
| **(d)**  **M1:** For using degrees of freedom to set up a model critical value  **A1:** For correct comparison and conclusion | | | |
| **(e)**  **A1ft:** For correct conclusion with supporting reason | | | |
| **Question** | **Scheme** | **Marks** | **AOs** |
| **2(a)** | –4 = 2 – 5E(*X*) | M1 | 3.1a |
| E(*X*) = 1.2 |  |  |
| –1×*c* + 0×*a* + 1×*a* + 2×*b* + 3×*c* = 1.2 | M1 | 1.1b |
| *a* + 2*b* + 2*c* = 1.2 |  |  |
| P(*Y*  – 3) = 0.45 gives P( 2 – 5*X*  – 3) = 0.45  i.e. P( *X*  1 ) = 0.45 | M1 | 2.1 |
| 2*a* + *c* = 0.45 |
| 2*a* + *b* + 2*c* = 1 | M1 | 1.1b |
| or | M1 | 1.1b |
| e.g. |
| *a* = 0.1 *b* = 0.3 *c* = 0.25 | A1  A1 | 1.1b  1.1b |
|  | **(7)** |  |
| **(b)** | Var(*Y*) = 75 – (–4)2 or 59 | M1 | 1.1a |
| [Var(*Y*) = 52Var(*X*) implies] Var(*X*) = 2.36 | A1 | 1.2 |
|  | **(2)** |  |
| **(c)** | P(*Y* > *X*) = P(2 – 5*X* > *X*) → P(*X* < ) | M1 | 3.1a |
| P(*X* < ) = *a* + *c* = 0.35 | A1ft | 1.1b |
|  | **(2)** |  |
| **(11 marks)** | | | |
| **Notes:** | | | |
| **(a)**  **M1:** For using given information to find an expression for E(*X*) i.e. use of E(*Y*) = 2 – 5E(*X*)  **M1:** For use of = ‘1.2’  **M1:** For use of P(*Y*  – 3 ) = 0.45 to set up the argument for solving by forming an equation  in *a* and *c*  **M1:** For use of  **M1:** For solving their 3 linear equations (matrix or elimination)  **A1:** For any 2 of *a*, *b* or *c* correct  **A1:** For all 3 correct values | | | |

|  |
| --- |
| **Question 2 notes continued:** |
| **Another method for part (a) is:**  **M1:** For using given information to find the probability distribution for *Y* leading to an  expression for E(*Y*)  **M1:** For use of = –4  **M1:** For use of P(*Y*  – 3 ) = 0.45 to set up the argument for solving by forming an equation  in *a* and *c*  **M1:** For use of  **M1:** For solving their 3 linear equations (matrix or elimination)  **A1:** For any 2 of *a*, *b* or *c* correct  **A1:** For all 3 correct values |
| **(b)**  **M1:**  For use of Var(*Y*) = E(*Y*2) – [E(*Y*)]2  (may be implied by a correct answer)  **A1:**  For use of Var(*aX*) = *a*2 Var(*X*) to reach 2.36 or exact equivalent |
| **(c)**  **M1:** For rearranging to the form P(*X* < *k*)  **A1ft:** 0.1’ + ‘025’ (provided their *a* and *c* and their *a* + *c* are all probabilities)  **Another method for part (c) is:**  **M1:** For comparing distribution of *X* with distribution of *Y* to identify *X* = –1 and *X* = 0  **A1ft:** ‘0.1’ + ‘025’ (provided their *a* and *c* and their *a* + *c* are all probabilities) |

|  |  |  |  |
| --- | --- | --- | --- |
| **Question** | **Scheme** | **Marks** | **AOs** |
| **3(a)** | *X* ~ Po(2.6) *Y* ~ Po(1.2) |  |  |
| P(each hire 2 in 1 hour)  = P(*X*=2) × P(*Y*=2) = 0.25104…× 0.21685… | M1 | 3.3 |
| = 0.05444… awrt **0.0544** | A1 | 1.1b |
|  | **(2)** |  |
| **(b)** | *W = X + Y → W ~* Po(3.8) | M1 | 3.4 |
| P(*W* = 3) = 0.20458…. awrt **0.205** | A1 | 1.1b |
|  | **(2)** |  |
| **(c)** | *T* ~ Po((2.6+1.2)×2) | M1 | 3.3 |
| P(*T* < 9) = 0.64819… awrt **0.648** | A1 | 1.1b |
|  | **(2)** |  |
| **(d)** | **(i)** Mean = *np* = **2.4** | B1 | 1.1b |
| **(ii)** Variance = *np*(1 – *p*) = 2.3904 awrt **2.39** | B1 | 1.1b |
|  | **(2)** |  |
| **(e)** | **(i)** [ *D* ~ Po(2.4) P(*D* 4) ]  = 0.9041… awrt **0.904** | B1 | 1.1b |
| **(ii)** Since *n* is large and *p* is small/mean is approximately equal to variance | B1 | 2.4 |
|  | **(2)** |  |
| **(10 marks)** | | | |
| **Notes:** | | | |
| **(a)**  **M1:** For P(*X*=2) × P(*Y*=2) from *X* ~ Po(2.6) and *Y* ~ Po(1.2) i.e. correct models (may be  implied by correct answer)  **A1: awrt 0.0544** | | | |
| **(b)**  **M1:** For combining Poisson distributions and use of Po(‘3.8’) (may be implied by correct  answer)  **A1:** **awrt 0.205** | | | |
| **(c)**  **M1:** For setting up a new model and attempting mean of Poisson distribution (may be implied  by correct answer)  **A1: awrt 0.648** | | | |
| **(d)(i)**  **B1:** For **2.4** | | | |
| **(d)(ii)**  **B1:** For **awrt 2.39** | | | |
| **(e)(i)**  **B1:** For **awrt 0.904** | | | |
| **(e)(ii)**  **B1:** For a correct explanation to support use of Poisson approximation in this case | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Question** | **Scheme** | **Marks** | **AOs** |
| **4(a)** | **(i)** P(*X* = 1) = 0.34523… awrt **0.345** | B1 | 1.1b |
| **(ii)** P(*X*  4) = 0.98575… awrt **0.986** | B1 | 1.1b |
|  | **(2)** |  |
| **(b)** | \* | B1\*cso | 1.1b |
|  | **(1)** |  |
| **(c)** | *r* = 40 × ‘0.34523…’ *s* = 40 × ‘1 – 0.986…’ | M1 | 3.4 |
| *r* = **13.81** *s* = **0.57** | A1ft | 1.1b |
|  | **(2)** |  |
| **(d)** | H0: The Poisson distribution is a suitable model  H1: The Poisson distribution is not a suitable model | B1 | 3.4 |
| [Cells are combined when expected frequencies < 5]  So combine the last 3 cells | M1 | 2.1 |
|  | M1 | 1.1b |
| awrt **1.1** | A1 | 1.1b |
| Degrees of freedom = 4 – 1 – 1 = 2 | B1 | 3.1b |
| (Do not reject H0 since 1.10 <). The number of mortgages approved each week follows a Poisson distribution | A1 | 3.5a |
|  | **(6)** |  |
| **(11 marks)** | | | |
| **Notes:** | | | |
| **(a)(i)**  **B1:** awrt 0.345 | | | |
| **(a)(ii)**  **B1:** awrt 0.986 | | | |
| **(b)**  **B1\*:** For a fully correct calculation leading to given answer with no errors seen | | | |
| **(c)**  **M1:** For attempt at *r* or *s* (may be implied by correct answers)  **A1ft:** For both values correct (follow through their answers to part (a)) | | | |
| **(d)**  **B1:** For both hypotheses correct (lambda should not be defined so correct use of the model)  **M1:** For understanding the need to combine cells before calculating the test statistic (may be  implied)  **M1:** For attempt to find the test statistic using  **A1:** awrt 1.1  **B1:** For realising that there are 2 degrees of freedom leading to a critical value  of  **A1:** Concluding that a Poisson model is suitable for the number of mortgages approved each  week | | | |