



ANSWER BOOK

Contingency tables

Further Statistics 1 · Year FM

6 questions · 20 marks

NAVY EDITION

SECTION A · Warm-up

- A1** H_0 is that the two variables are independent, so there is no association between them [2]
Each expected frequency is row total $\times$ column total $\div$ grand total.
- A2** $(4 - 1)(5 - 1) = 12$. No further subtraction is made for estimated parameters: the row and column totals have already absorbed them.

SECTION B · Standard

- B1** Row totals are 50 and 50; column totals 40, 40 and 20; grand total 100. Expected values: $50 \times 40/100 = 20$ for the first two columns of each row, and $50 \times 20/100 = 10$ for the third. Equal row totals make both rows expect the same split here.
- B2** The contributions are $25/20$, $25/20$, 0, $25/20$, $25/20$ and 0, totalling 5.0. With $(2 - 1)(3 - 1) = 2$ degrees of freedom and a critical value of 5.991, we have $5.0 < 5.991$: do not reject H_0 . There is insufficient evidence at the 5% level of an association, though the result is close to the boundary.
- B3** Once both row totals and both column totals are fixed, choosing any single cell [4]
determines the other three by subtraction. Exactly one cell is free to vary, and $(2 - 1)(2 - 1) = 1$ records that.

SECTION C · Stretch

- C1** It may be concluded that there is evidence of an association between the two [4]
variables: the observed pattern is unlikely if they were independent. It may not be
concluded that one variable causes the other, nor that the association is large: the test
measures evidence against independence, not the strength or direction of any effect,
and a large sample can make a slight association significant.