inkMaths

ANSWER BOOK

Testing variances: chi-squared and the F-distribution

Further Statistics 2 · Year FM

7 questions · 25 marks

NAVY EDITION

inkphysics.co.uk/maths

SECTION A · Warm-up

A1 The statistic is $(n - 1)S^2/\sigma^2$, where S^2 is the sample variance and σ^2 the value claimed under the null hypothesis. It follows a chi-squared distribution on $n - 1$ degrees of freedom, provided the population is normal. B1 for the statistic, B1 for the chi-squared distribution on $n - 1$ degrees of freedom. [2]

A2 H_0: $\sigma^2 = 4$; H_1: $\sigma^2 > 4$, one-tailed at 5% with 24 degrees of freedom. The statistic is $24 \times 6.4/4 =$ **38.4**. Since $38.4 > 36.415$, **reject** H_0: there is evidence at the 5% level that the variance exceeds 4. B1 for the hypotheses, M1 for the statistic, A1 for 38.4 with the conclusion in context. [3]

SECTION B · Standard

B1 The numerator is $(n - 1)s^2 = 24 \times 6.4 =$ **153.6**. Dividing by the larger chi-squared value gives the lower limit, $153.6/39.364 =$ **3.90**, and dividing by the smaller gives the upper limit, $153.6/12.401 =$ **12.39**. The interval is **(3.90,** 12.39). M1 for $(n - 1)s^2$, A1 for 153.6, M1 for dividing by both chi-squared values, A1 for the interval. The two ends look swapped because the chi-squared value sits in the denominator, and pairing each limit with the wrong tail is the standard slip. [4]

B2 H_0: $\sigma_1^2 = \sigma_2^2$; H_1: $\sigma_1^2 > \sigma_2^2$. The alternative names the first population, so the first sample variance goes on top: $F = 24.5/8.2 =$ **2.99**, on **12** and 8 degrees of freedom, numerator first. [4]

Since $2.99 < 3.284$, **do** not reject H_0. There is insufficient evidence at the 5% level that the first population is more variable, even though one sample variance is about three times the other. B1 for the hypotheses, M1 for $F = 24.5/8.2$, A1 for 2.99 on 12 and 8 degrees of freedom, A1 for not rejecting H_0 in context.

B3 For a two-tailed test, put the **larger** sample variance on top, so that F is at least 1 and only the printed upper tail is needed. For a one-tailed test, the alternative hypothesis decides it: the population claimed to be more variable supplies the numerator. Either way the degrees of freedom must follow the variances, numerator first, since the F distribution is not symmetric in its two parameters. B1 for the larger variance on top in a two-tailed test, B1 for the degrees of freedom following the variances. [2]

B4 $H_0: \sigma^2 = 5$; $H_1: \sigma^2 \neq 5$, with 15 degrees of freedom and 2.5% in each tail. The statistic is $15 \times 2.5/5 =$ **7.5**. Since $6.262 < 7.5 < 27.488$ the value lies between the two critical values, so **do** not reject H_0. [4]

B1 for the hypotheses, M1 for the statistic, A1 for 7.5, A1 for not rejecting H_0 in context. Note how far apart the critical values are. The chi-squared distribution is strongly skewed, so the acceptance region is nothing like symmetric about the mean of 15.

SECTION C · Stretch

C1 $H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 \neq \sigma_2^2$, two-tailed. [6]

A two-tailed test at 10% puts **5%** in each tail, so the 5% critical value is the one to use. Comparing with the 10% value would double the size of the test.

The alternative names no direction, so put the larger sample variance on top: $F = 12.6/4.2 =$ **3.00**. That variance came from the sample of 16, so the numerator has 15 degrees of freedom and the denominator has 9. The degrees of freedom follow the variances, in that order.

Since $3.00 < 3.006$, **do** not reject H_0. There is insufficient evidence at the 10% level that the population variances differ. It is as close as a decision gets, and quoting the statistic to three significant figures is what keeps the comparison honest.

Taking the ratio the other way up gives $4.2/12.6 = 0.333$ on 9 and 15 degrees of freedom. A value below 1 must be judged against the **lower** 5% point of F on 9 and 15, and the tables print only upper points, so an extra step is needed: the lower point is the reciprocal of the printed 3.006. Putting the larger variance on top avoids that step entirely. B1 for the hypotheses, B1 for 5% in each tail, M1 for $F = 12.6/4.2$, A1 for 3.00 on 15 and 9 degrees of freedom, A1 for not rejecting H_0 in context, B1 for the effect of inverting the ratio.