inkMaths

ANSWER BOOK

Hypothesis testing with the binomial

Statistics · Year 12–13

7 questions · 29 marks

NAVY EDITION

inkphysics.co.uk/maths/ · youtube.com/@InkPhysics

SECTION A · Warm-up

A1 The null hypothesis H_0 states the assumed value of the population parameter, the value taken as true while the test runs. The alternative H_1 states the change being tested for. The significance level is the probability threshold, so evidence rarer than that threshold under H_0 counts as grounds for rejection. B1 B1 B1, one for each of the three definitions. Both hypotheses must name a parameter, never an observed count. [3]

A2 H_0: $p = 1/6$ and H_1: $p > 1/6$, where p is the probability that a single roll gives a six. Under H_0, $X \sim B(30, 1/6)$. B1 B1 for the two hypotheses and B1 for the distribution. 'Biased towards sixes' points one way only, so the alternative is one-tailed with a strict inequality. [3]

SECTION B · Standard

B1 H_0: $p = 0.3$ and H_1: $p > 0.3$, where p is the probability of red on one spin. Under H_0, $X \sim B(25, 0.3)$. The probability of a result at least as extreme as the one observed is $P(X \geq 12) = 1 - P(X \leq 11) = 0.0442$. Since $0.0442 < 0.05$, reject H_0. There is significant evidence at the 5% level that the spinner lands on red more often than claimed. B1 hypotheses, M1 the binomial model, M1 the correct tail, A1 0.0442, A1 comparison and conclusion in context. Working out $P(X \leq 12)$ on the calculator instead of $1 - P(X \leq 11)$ gives 0.9827 and a conclusion pointing the wrong way. [5]

B2 Work down from the top of the distribution. $P(X \geq 12) = 0.0565$, which is bigger than 0.05, so 12 cannot be in the region. $P(X \geq 13) = 0.0210$, which is smaller, so 13 can. The critical region is $X \geq 13$, and the actual significance level is 0.0210, or 2.10%. M1 for attempting a tail, A1 for 0.0565, A1 for the region, B1 for the actual level. Because X is discrete the advertised 5% is almost never achievable exactly, and the actual level is the true tail probability of whatever region you end up with. [4]

B3 12 lies outside the critical region, so do not reject H_0. There is insufficient evidence at the 5% level that p has risen above 0.4. That is a failure to disprove, not a proof. The data are consistent with $p = 0.4$ and equally consistent with $p = 0.45$ or $p = 0.5$, none of which the test has ruled out. M1 for comparing with the region, A1 for the conclusion in context, B1 for the explanation. A hypothesis test never proves the null, and writing 'so $p = 0.4$' throws away the last mark. [3]

SECTION C · Stretch

C1 Under H_0, $X \sim B(30, 0.25)$, and a two-tailed 5% test puts at most 2.5% in each tail. Lower end: $P(X \le 2) = 0.0106$, which is under 0.025, while $P(X \le 3) = 0.0374$, which is over, so the lower region is $X \le 2$. Upper end: $P(X \ge 13) = 0.0216$, under 0.025, while $P(X \ge 12) = 0.0507$, over, so the upper region is $X \ge 13$. The critical region is $X \le 2$ or $X \ge 13$, and the actual significance level is $0.0106 + 0.0216 = 0.0322$, or 3.22%. B1 for splitting the 5% into two 2.5% halves, M1 A1 for the lower tail, M1 A1 for the upper tail, A1 for the actual level. Two errors dominate here. Comparing each tail with 0.05 rather than 0.025 makes the region far too big, and quoting 5% as the actual level ignores the whole point of the last part. [6]

C2 Under H_0, $X \sim B(30, 0.3)$, and the alternative points downwards so the region sits in the lower tail. $P(X \le 3) = 0.00932$, which is under 0.01, while $P(X \le 4) = 0.0302$, which is over. The critical region is therefore $X \le 3$. A Type I error is rejecting H_0 when it is true, which happens exactly when X lands in the critical region while p really is 0.3, so its probability is 0.00932. M1 for working in the lower tail, A1 for 0.00932, A1 for 0.0302, A1 for the region, B1 for the Type I probability. The probability of a Type I error and the actual significance level are the same number, and questions ask for it under either name. [5]