InkMaths

ANSWER BOOK

Hypothesis tests for Poisson and geometric models

Further Statistics 1 · Year FM

7 questions · 25 marks

NAVY EDITION

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

SECTION A · Warm-up

A1 H_0: $\lambda = 3$ against H_1: $\lambda > 3$, a one-tailed test. B1 B1 for the two hypotheses. The hypotheses name the population parameter, never the observed count. [2]

A2 H_0: $\lambda = 3$ against H_1: $\lambda > 3$, one-tailed at 5%. Under H_0, $P(X \geq 8) = 1 - P(X \leq 7) =$ **0.0119**. Since 0.0119 < 0.05, **reject** H_0. There is evidence at the 5% level that the fault rate has risen. M1 for the correct upper tail, A1 for 0.0119, A1 for the comparison and conclusion in context. [3]

SECTION B · Standard

B1 H_0: $\lambda = 10$, H_1: $\lambda < 10$, one-tailed. The observed value is low, so use the lower tail: $P(X \leq 4) = 0.0293$. Since 0.0293 < 0.05, reject H_0: there is evidence at the 5% level that the accident rate has fallen since the campaign. B1 for the hypotheses, M1 for using the lower tail, A1 for 0.0293, A1 for the conclusion in context. [4]

B2 H_0: $p = 0.4$ against H_1: $p < 0.4$. A small p produces a long wait, so the evidence sits in the upper tail of X: $P(X \geq 8) = 0.6^7 = 0.0280$. Since 0.0280 < 0.05, reject H_0: there is evidence that the success probability is below 0.4. B1 for the hypotheses, M1 for the upper tail, A1 for 0.0280, A1 for the conclusion in context. [4]

B3 The 5% is split between the two tails, so each carries only 2.5%. The observed tail probability is compared with 0.025 rather than 0.05, which makes rejection harder and can reverse the conclusion of an otherwise identical calculation. B1 for the split into two tails of 2.5% each, B1 for comparing with 0.025 rather than 0.05. [2]

B4 $P(X \geq 8) = 1 - P(X \leq 7) = 1 - 0.9489 =$ **0.0511**, which exceeds 0.05, so 8 is not extreme enough to reject. $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.9786 =$ **0.0214**, which is below 0.05.
The critical region is therefore $\mathbf{X} \geq 9$, with actual significance level **0.0214**. M1 for forming an upper tail from the tabulated values, A1 for 0.0511 ruling out a count of 8, A1 for the critical region, A1 for the actual significance level. Take the tail one place above the tabulated value: reading $P(X \geq 8)$ as $1 - P(X \leq 8)$ is the standard error and shifts the whole region. [4]

SECTION C · Stretch

C1 H_0: $\lambda = 2.5$ per minute against H_1: $\lambda > 2.5$, one-tailed. [6]

Scale the parameter to the interval actually observed. Over 4 minutes the mean is $4 \times 2.5 = \mathbf{10}$, so under H_0 the count X follows Po(10). Testing an 18 against a mean of 2.5 is the error that wrecks this question.

Work down the upper tail. $P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9165 = 0.0835$, which is above 5%, so 15 is not extreme enough. $P(X \geq 16) = 1 - 0.9513 = 0.0487$, which is below 5%.

The critical region is $\mathbf{X \geq 16}$, and the actual significance level is **0.0487**, or 4.87%. It undershoots the nominal 5% because X takes only whole values, so the tail cannot be trimmed to an exact 5%.

The observed value 18 lies in the critical region, so **reject** H_0. There is evidence at the 5% level that the rate at which cars pass the checkpoint has increased. Equivalently $P(X \geq 18) = 0.0143 < 0.05$.

B1 for the hypotheses, M1 for scaling the mean to the 4-minute interval, M1 for working down the upper tail, A1 for the critical region, A1 for the actual significance level, A1 for the conclusion in context.