2.4 Further Statistics: Hypothesis Testing, Type I and Type II Errors: 1-hour Pack

AQA 7367 · Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

Hypothesis Testing for a Poisson or Binomial Parameter

Revision Library

For a hypothesis test on a Poisson mean lambda or a binomial probability p, a null hypothesis H0 states a fixed value of the parameter, and an alternative hypothesis H1 states that the parameter has increased, has decreased (a one-tailed test), or is simply different (a two-tailed test). A test statistic X, assumed to follow the distribution stated in H0, is compared against a critical region: the set of values of X for which H0 is rejected in favour of H1.

Because X is discrete, the probability of X falling in a proposed critical region can rarely be made to equal a chosen significance level exactly. The usual convention is to choose the largest critical region whose probability does not exceed the significance level requested. This probability is called the actual significance level of the test, and it is also equal to P(Type I error), the probability of wrongly rejecting a true H0.

A Type II error occurs when H0 is not rejected even though H1 is actually true. Its probability depends on the true value of the parameter, and is found by calculating the probability that X falls outside the critical region, using the true (alternative) value of the parameter in place of the value stated in H0. The power of a test is defined as 1 - P(Type II error): the probability of correctly rejecting a false H0. Power increases as the true parameter value moves further from the value stated in H0, and is also affected by the significance level chosen and the sample size used.

1 The number of faults in a length of fabric is modelled by a Poisson distribution with mean λ. A test of H0: $\lambda = 5$ against H1: $\lambda > 5$ is carried out at the 5% significance level, using a single observation X of the number of events.

(a) Find the critical region for the test. **(3)**

(b) State the probability of a Type I error for this test, to 3 significant figures. **(2)**

(c) Given that the true value of λ is 6, find the probability of a Type II error, to 3 significant figures. **(2)**

(Total for Question 1 is 7 marks)

© Revision Library | revisionlibrary.co.uk

2 The random variable $X \sim B(25, p)$. A test of H0: $p = 0.25$ against H1: $p < 0.25$ is carried out at the 10% significance level.

(a) Find the critical region for the test. **(2)**

(b) State the probability of a Type I error, to 3 significant figures. **(2)**

(c) Find the probability of a Type II error when the true value of p is 0.1, to 3 significant figures. **(2)**

(Total for Question 2 is 6 marks)

---

3 The number of customers arriving per minute is modelled by a Poisson distribution with mean λ. A test of H0: $\lambda = 8$ against H1: $\lambda < 8$ is carried out at the 10% significance level, using a single observation X of the number of events.

(a) Find the critical region for the test. **(3)**

(b) State the probability of a Type I error for this test, to 3 significant figures. **(2)**

(c) Given that the true value of λ is 3, find the probability of a Type II error, to 3 significant figures. **(2)**

(Total for Question 3 is 7 marks)

---

4 The random variable $X \sim B(12, p)$. A test of H0: $p = 0.2$ against H1: $p > 0.2$ is carried out at the 10% significance level.

(a) Find the critical region for the test. **(2)**

(b) State the probability of a Type I error, to 3 significant figures. **(2)**

(c) Find the probability of a Type II error when the true value of p is 0.7, to 3 significant figures. **(2)**

(Total for Question 4 is 6 marks)

© Revision Library | revisionlibrary.co.uk

5 A fair coin-like device is claimed to land on 'success' with probability 0.1. A tester believes the probability of success is actually lower. Let X be the number of trials needed to get the first success, so under the claim X ~ Geo(0.1). The first success occurs on trial 12.

(a) State suitable hypotheses for the test, in terms of p. **(2)**

(b) Find P(X ≥ 12) assuming H0 is true, and test the tester's belief at the 10% level. **(3)**

(c) Find the critical region for the test at the 10% level, and the probability of a Type II error if the true value of p is 0.05, to 3 significant figures. **(4)**

(Total for Question 5 is 9 marks)

---

6 The number of minor faults found on newly manufactured circuit boards from a particular production line is modelled, under normal conditions, by X ~ Po(6). Quality control engineers want to test, at the 10% significance level, whether the mean number of faults per board has changed from 6 (a two-tailed test).

(a) State suitable hypotheses, then find the critical region for this two-tailed test and state the actual significance level. **(6)**

(b) A randomly chosen board is found to have 10 faults. State, with a reason, the conclusion of the test. **(2)**

(c) Suppose that a fault in the manufacturing process has in fact increased the true mean number of faults per board to 9. Find the probability that the test results in a Type II error. **(3)**

(d) Hence state the power of the test in this case. **(1)**

(e) Comment on the suitability of this test for detecting the increase in mean described in part (c). **(1)**

(Total for Question 6 is 13 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 2.4 Further Statistics: Hypothesis Testing, Type I and Type II Errors: 1-hour Pack

Question 1

(a) uses Po(5) probabilities in the appropriate tail **M1**

(a) $P(X \geq 10) = 0.0318 \leq 0.05$ **A1**

(a) critical region $X \geq 10$ **A1**

(a) Answer: critical region $X \geq 10$

(b) the probability of a Type I error = P(X in the critical region | H0 true) **M1**

(b) 0.0318 **A1**

(b) Answer: 0.0318

(c) P(X not in the critical region | $\lambda = 6$) **M1**

(c) 0.916 **A1**

(c) Answer: 0.916

Question 2

(a) uses B(25, 0.25) cumulative probabilities **M1**

(a) critical region $X \leq 3$ **A1**

(a) Answer: critical region $X \leq 3$

(b) P(X in critical region | $p = 0.25$) **M1**

(b) 0.0962 **A1**

(b) Answer: 0.0962

(c) P(X not in critical region | $p = 0.1$) **M1**

(c) 0.236 **A1**

(c) Answer: 0.236

Question 3

(a) uses Po(8) probabilities in the appropriate tail **M1**

(a) $P(X \leq 4) = 0.0996 \leq 0.1$ **A1**

(a) critical region $X \leq 4$ **A1**

(a) Answer: critical region $X \leq 4$

(b) the probability of a Type I error = P(X in the critical region | H0 true) **M1**

(b) 0.0996 **A1**

(b) Answer: 0.0996

(c) P(X not in the critical region | $\lambda = 3$) **M1**

(c) 0.185 **A1**

(c) Answer: 0.185

Question 4

(a) uses B(12, 0.2) cumulative probabilities **M1**

(a) critical region $X \geq 5$ **A1**

(a) Answer: critical region $X \geq 5$

(b) P(X in critical region | $p = 0.2$) **M1**

(b) 0.0726 **A1**

(b) Answer: 0.0726

(c) P(X not in critical region | $p = 0.7$) **M1**

(c) 0.00949 **A1**

(c) Answer: 0.00949

Question 5

(a) H0: $p = 0.1$ **B1**

(a) H1: $p < 0.1$ **B1**

(a) Answer: H0: $p = 0.1$, H1: $p < 0.1$

(b) $P(X \geq 12) = (1 - 0.1)^{12 - 1}$ **M1**

(b) 0.314 **A1**

(b) do not reject H0 **A1**

(b) Answer: 0.314; do not reject H0

(c) finds the smallest d with $(1 - 0.1)^{d - 1} \leq 0.1$ **M1**

(c) critical region $X \geq 23$ **A1**

(c) $P(X \leq 22 \mid p = 0.05) = 1 - (1 - 0.05)^{22}$ **M1**

(c) 0.676 **A1**

(c) Answer: $X \geq 23$; Type II error 0.676

Question 6

(a) states H0: $\lambda = 6$, H1: $\lambda \neq 6$, and searches the lower tail comparing $P(X \leq 1)$ and $P(X \leq 2)$ with 0.05 **M1**

(a) lower critical region $X \leq 1$ (i.e. {0, 1}), actual lower-tail level = awrt 0.0174 **A1**

(a) searches the upper tail comparing $P(X \geq 10)$ and $P(X \geq 11)$ with 0.05 **M1**

(a) upper critical region $X \geq 11$ (i.e. {11, 12, 13, ...}), actual upper-tail level = awrt 0.0426 **A1**

(a) combined critical region stated as {0, 1} union {11, 12, 13, ...} **A1**

(a) total actual significance level = 0.0174 + 0.0426 = awrt 0.0600 **A1**

(a) Answer: Critical region: {0, 1} union {11, 12, 13, ...}. Total actual significance level = awrt 0.0600.

(b) 10 does not lie in the critical region {0, 1} union {11, 12, ...} found in part (a) (ft) **B1**

(b) do not reject H0; there is insufficient evidence, at the 10% level, that the mean number of faults per board has changed (ft, in context) **B1**

(b) Answer: 10 is not in the critical region, so H0 is not rejected; there is insufficient evidence at the 10% level that the mean has changed.

(c) identifies the acceptance (non-rejection) region as $2 \leq X \leq 10$, using the true distribution Po(9) **M1**

(c) uses $P(2 \leq X \leq 10 \mid \lambda = 9) = P(X \leq 10 \mid \lambda = 9) - P(X \leq 1 \mid \lambda = 9)$ **M1**

(c) awrt 0.705 **A1**

(c) Answer: P(Type II error) = awrt 0.705

(d) power = 1 - 0.705 = awrt 0.295 (ft from part (c)) **B1**

(d) Answer: power = awrt 0.295

(e) valid comment referencing the low power value, e.g. the power is only awrt 0.295, so the test would fail to detect this particular increase in mean on roughly 70% of occasions; the test is not very sensitive to a change of this size (oe) **B1**

(e) Answer: The power is only about 0.295, so the test would fail to detect an increase in mean to 9 on around 70% of occasions; it is not a very sensitive test for a change of this size.

© Revision Library | revisionlibrary.co.uk