KS5 "Full Coverage": Hypothesis Testing

This worksheet is designed to cover one question of each type seen in past papers, for each A Level topic. This worksheet was automatically generated by the DrFrostMaths Homework Platform: students can practice this set of questions interactively by going to www.drfrostmaths.com, logging on, *Practise → Past Papers* (or *Library → Past Papers* for teachers), and using the 'Revision' tab.

Question 1

Categorisation: Conduct a one-tailed hypothesis testing on the value p of a Binomial distribution.

[Edexcel AS Specimen Papers P2 Q2b Edited]

Past records from a large supermarket show that 25% of people who buy eggs, buy organic eggs. On one particular day, a random sample of 40 people is taken from those that had bought eggs and 16 people are found to have bought organic eggs.

Test, at the 1% significance level, whether or not the proportion, p , of people who bought organic eggs that day had increased. State your hypotheses clearly.

....

Question 2
Categorisation: As above, but at a different significance level.

[Edexcel AS Specimen Papers P2 Q2c Edited]

Past records from a large supermarket show that 25% of people who buy eggs, buy organic eggs. On one particular day, a random sample of 40 people is taken from those that had bought eggs and 16 people are found to have bought organic eggs.

Test, at the 5% significance level, whether or not the proportion, p , of people who bought organic eggs that day had increased. State your hypotheses clearly.

……………………

Question 3
Categorisation: State the hypotheses required for a one or two-tailed test.

[Edexcel AS SAM P2 Q5b Edited]

Past records suggest that 30% of customers who buy baked beans from a large supermarket buy them in single tins. A new manager suspects that there has been a change in the proportion of customers who buy baked beans in single tins. A random sample of 20 customers who had bought baked beans was taken.

Write down the hypotheses that should be used to test the manager's suspicion.

Question 4

Categorisation: Determine the critical region for a two-tailed Binomial test.

[Edexcel AS SAM P2 Q5c Edited]

Past records suggest that 30% of customers who buy baked beans from a large supermarket buy them in single tins. A new manager suspects that there has been a change in the proportion of customers who buy baked beans in single tins. A random sample of 20 customers who had bought baked beans was taken.

Using a 10% level of significance, find the critical region for a two-tailed test to answer the manager's suspicion. You should state the probability of rejection in each tail, which should be less than 0.05

Question 5

Categorisation: Calculate the actual significance level of a test.

[Edexcel AS SAM P2 Q5d Edited]

(Continued from above) Find the actual significance level of a test based on this critical region.

……………………… %

Question 6

Categorisation: Comment on an observed value using a given/calculated critical region.

[Edexcel AS SAM P2 Q5e Edited]

(Continued from above) One afternoon the manager observes that 12 of the 20 customers who bought baked beans, bought their beans in single tins.

Comment on the manager's suspicion in the light of this observation.

Question 7

Categorisation: Conduct a hypothesis test on the correlation ρ between two variables.

[Edexcel A2 Specimen Papers P3 Q2b]

A researcher believes that there is a linear relationship between daily mean temperature and daily total rainfall. The 7 places in the northern hemisphere from the large data set are used. The mean of the daily mean temperatures, ts mm, for the month of July in 2015 are shown on the scatter diagram below.
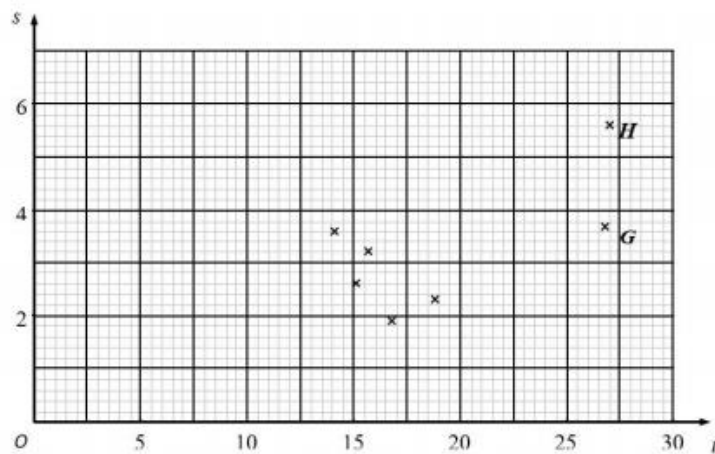
The researcher calculated the product moment correlation coefficient for the 7 places and obtained $r = 0.658$

Stating your hypotheses clearly, test at the 10% level of significance, whether or not the product moment correlation coefficient for the population is greater than zero.

....

Question 8

Categorisation: As above.

[Edexcel A2 SAM P3 Q2c Edited] A meteorologist believes that there is a relationship between the daily mean windspeed, w kn, and the daily mean temperature, t

t	13.3	16.2	15.7	16.6	16.3	16.4	19.3	17.1	13.2
w	7	11	8	11	13	8	15	10	11

The meteorologist calculated the product moment correlation coefficient for the 9 days and obtained $r = 0.609$

Stating your hypotheses clearly test, at the 5% significance level, whether or not the product moment correlation coefficient for the population is greater than zero.

....

Question 9

Categorisation: Conduct a hypothesis test on the mean μ.

[Edexcel A2 SAM P3 Q3c Edited] A machine cuts strips of metal of length X cm, where X is normally distributed with standard deviation 0.6 cm. A random sample of 15 strips cut by this machine was found to have a mean length of 50.4 cm

Stating your hypotheses clearly and using a 1% level of significance, test whether or not the mean length of all the strips, cut by the machine, is greater than 50.1 cm

....

Question 10
Categorisation: As above.

[Edexcel A2 Specimen Papers P3 Q1d Edited] Kaff coffee is sold in packets. A seller measures the masses of the contents of a random sample of 90 packets of *Kaff coffee* from her stock. The results are shown in the table below.

Mass w (g)	Midpoint y (g)	Frequency (f)
$240 \leq w < 245$	242.5	8
$245 \leq w < 248$	246.5	15
$248 \leq w < 252$	250	35
$252 \leq w < 255$	253.5	23
$255 \leq w < 260$	257.5	9

(You may use $\sum fy^2 = 5\,644\,171.75$)

A histogram is drawn and the class $245 \leq w < 248$ is represented by a rectangle of width 1.2 cm and height 10 cm.

The seller claims that the mean mass of the contents of the packets is more than the stated mass. Given that the stated mass of the contents of a packet of Kaff coffee is 250 g and the actual standard deviation of the contents of a packet of Kaff coffee is 4 g, test, using a 5% level of significance, whether or not the seller's claim is justified.
State your hypotheses clearly. *(You may assume that the mass of the contents of a packet is normally distributed.)*

....

Question 11
Categorisation: Define key terms such as 'test statistic'.

[Edexcel Statistics Year 2 Unit Test: Regression & Correlation Q7a]
State the definition of a test statistic.

....

Question 12

Categorisation: Understand that the p-value is just the probability of seeing the observed outcome 'or more extreme'.
Calculate a p-value or conduct a hypothesis test when the p-value is already given (by comparing against the significance threshold).

[Edexcel Statistics Year 2 Unit Test: Regression & Correlation Q7b]
   To investigate if there is a correlation between daily mean temperature (°C) and daily mean pressure (hPa) the location Hurn 2015 was randomly selected from:

Camborne 2015	Camborne 1987
Hurn 2015	Hurn 1987
Leuchars 2015	Leuchars 1987
Leeming 2015	Leeming 1987
Heathrow 2015	Heathrow 1987

The product moment correlation coefficient between daily mean temperature and daily mean pressure for these data is −0.258 with a p-value of 0.001. Use a 5% significance level to test whether or not there is evidence of a correlation between the daily mean temperature and daily mean pressure.

Answers

Question 1

$H_0: p = 0.25$ $H_1: p > 0.25$ (both correct in terms of p or p)		B1	2.5
$Y \sim B(40, 0.25)$		M1	3.3
Method 1	**Method 2**		
$P(Y \geq 16) = 1 - P(Y \pounds 15)$	$P(Y \geq 17) = 0.0116$	M1	1.1b
$= 1 - 0.9738$	$P(Y \geq 18) = 0.0047$		
$= 0.0262$	CR $Y \geq 18$	A1	1.1b
$0.0262 > 0.01$ $16 < 18$ or 16 is not in the critical region or 16 is not significant, accept H_0. There is no significant evidence that the proportion of people who bought organic eggs has increased		A1cso	2.2b

Question 2

$H_0: p = 0.25$ $H_1: p > 0.25$ (both correct in terms of p or p)		B1	2.5
$Y \sim B(40, 0.25)$		M1	3.3
Method 1	**Method 2**		
$P(Y \geq 16) = 1 - P(Y \pounds 15)$	$P(Y \geq 17) = 0.0116$	M1	1.1b
$= 1 - 0.9738$	$P(Y \geq 18) = 0.0047$		
$= 0.0262$	CR $Y \geq 18$	A1	1.1b

Question 3

$H_0 : p = 0.3$ $H_1 : p \neq 0.3$ (Both correct in terms of p or π)	B1

Question 4

$[Y \sim B(20, 0.3)]$ sight of $P(Y \leq 2) = 0.0355$ or $P(Y \leq 9) = 0.9520$	M1
Critical region is $\{Y \leq 2\}$ or (o.e.)	A1
$\{Y \geq 10\}$ (o.e.)	A1

Question 5

8.35 %

Question 6

(Assuming that the 20 customers represent a random sample then) 12 is in the CR so the manager's suspicion is supported	B1ft

Question 7

$H_0 : \rho = 0$ $H_1 : \rho > 0$	B1	2.5
Critical value 0.5509	M1	1.1a
Reject H_0		
There is evidence that pmcc is greater than zero	A1	2.2b

Question 8

H_0: $\rho = 0$ H_1: $\rho > 0$	B1
Critical value 0.5822	M1
Reject H_0	
There is evidence that the product moment correlation coefficient is greater than 0	A1

Question 9

$H_0 : \mu = 50.1$ $H_1 : \mu > 50.1$	B1
$\bar{X} \sim N\left(50.1, \dfrac{0.6^2}{15}\right)$ and $\bar{X} > 50.4$	M1
$P(\bar{X} > 50.4) = 0.0264$	A1
$p = 0.0264 > 0.01$ or $z = 1.936... < 2.3263$ and not significant	A1
There is insufficient evidence that the **mean length** of strips is **greater than 50.1**	A1

Question 10

$H_0 : \mu = 250$ $H_1 : \mu > 250$	B1	2.5
$\bar{X} \sim N\left(250, \dfrac{4^2}{90}\right)$ and $\bar{X} > 250.4$	M1	3.3
$P(\bar{X} > 250.4) = 0.171...$	A1	3.4
$0.171 > 0.05$ or $z = 0.9486... < 1.6449$	A1	1.1b
There is insufficient evidence that the mean weight of coffee is greater than 250 g, or there is no evidence to support the sellers claim.	A1	2.2b

Question 11

A statistic that is calculated from sample data in order to test a hypothesis about a population.	**B1**

Question 12

$H_0: = 0, H_1: \neq 0$	**B1**
p-value < 0.05	**M1**
There is evidence to reject H_0	
There is evidence (at 5% level) of a correlation between the daily mean temperature and daily mean pressure.	**A1**