HYPOTHESIS TESTING

INTRODUCTION

BINOMIAL DISTRIBUTION HYPOTHESIS INTRODUCTION

Question 1

In a craft activity in a primary school, kids use beads which are kept in a bag. The bag contains a large number of beads of different colours. The beads are not replaced into the bag at the end of the activity.

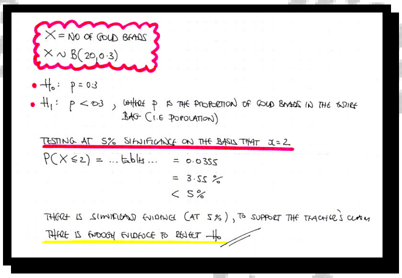
It is known that $\frac{3}{10}$ of the beads are coloured gold.

The teacher claims that children use more gold beads during the activity and checks a random sample of 20 beads out of the bag, after the end of the activity.

She finds just two gold beads in the sample.

Test, at the 5% level of significance, whether or not there is evidence to support the teacher's claim.

$\boxed{\text{significant evidence, } 3.55\% < 5\%}$



X = NO OF GOLD BEADS

$X \sim B(20, 0.3)$

- $H_0: p = 0.3$
- $H_1: p < 0.3$, WHERE p IS THE PROPORTION OF GOLD BEADS IN THE ENTIRE BAG (I.E POPULATION)

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 2$

$P(X \leq 2) = \ldots \text{tables} \ldots = 0.0355$

$= 3.55\%$

$< 5\%$

THERE IS SIGNIFICANT EVIDENCE (AT 5%), TO SUPPORT THE TEACHER'S CLAIM

THERE IS ENOUGH EVIDENCE TO REJECT H_0

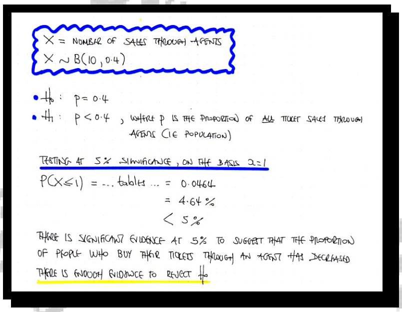
Question 2

A theatre company finds from its records that 40% of its customers book their tickets through agents. The company redesigns its website and then carries out a survey with 10 randomly chosen customers.

The result of the survey is that 1 of these 10 customers booked their tickets through an agent.

Test, at the 5% level of significance, whether the percentage of customers who book their tickets through an agent has decreased.

$\boxed{\text{significant, } 4.64\% < 5\%}$



X = NUMBER OF SALES THROUGH AGENTS
$X \sim B(10, 0.4)$

- $H_0: \ p = 0.4$
- $H_1: \ p < 0.4$, WHERE p IS THE PROPORTION OF ALL TICKET SALES THROUGH AGENTS (IE POPULATION)

TESTING AT 5% SIGNIFICANCE, ON THE BASIS $x=1$

$P(X \leqslant 1) = \dots \text{tables} \dots = 0.0464$
$= 4.64\%$
$< 5\%$

THERE IS SIGNIFICANT EVIDENCE AT 5% TO SUGGEST THAT THE PROPORTION OF PEOPLE WHO BUY THEIR TICKETS THROUGH AN AGENT HAS DECREASED
THERE IS ENOUGH EVIDENCE TO REJECT H_0

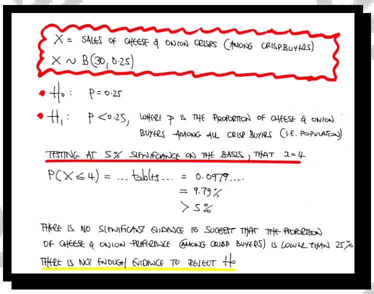
Question 3

The owner of a corner shop believes that 25% of the customers who buy crisps will buy the "cheese and onion" variety.

He finds that in the last 30 customers who bought crisps, only 4 customers bought the "cheese and onion" variety.

Test, at the 5% level of significance, whether there is evidence to suggest that the proportion of customers who choose the "cheese and onion" variety is lower than 25%.

$\boxed{\text{not significant evidence, } 9.79\% > 5\%}$



X = SALES OF CHEESE & ONION CRISPS (AMONG CRISP BUYERS)
$X \sim B(30, 0.25)$

- $H_0: \ p = 0.25$
- $H_1: \ p < 0.25$, WHERE p IS THE PROPORTION OF CHEESE & ONION BUYERS AMONG ALL CRISP BUYERS (I.E. POPULATION)

TESTING AT 5% SIGNIFICANCE ON THE BASIS, THAT $x=4$

$P(X \leqslant 4) = \dots \text{tables} \dots = 0.0979\dots$
$= 9.79\%$
$> 5\%$

THERE IS NO SIGNIFICANT EVIDENCE TO SUGGEST THAT THE PROPORTION OF CHEESE & ONION PREFERENCE (AMONG CRISP BUYERS) IS LOWER THAN 25%
THERE IS NOT ENOUGH EVIDENCE TO REJECT H_0

Question 4

The owner of a "Fish and Chips" shop believes that 35% of the customers who buy chips like vinegar with their chips.

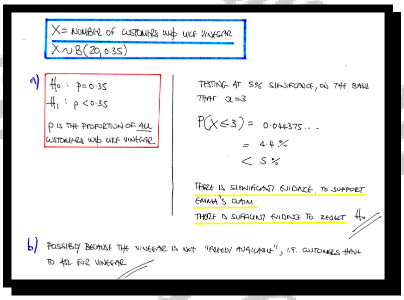
His wife Emma, claims that this proportion is lower and decides to test it by checking whether the next 20 customers who buy chips will ask for vinegar.

Only 3 customers in Emma's sample asked for vinegar.

a) Test, at the 5% level of significance, whether or not there is evidence to support Emma's claim.

b) State a reason as to why this method of testing might produce a biased result.

$\boxed{\text{significant evidence, } 4.44\% < 5\%}$, $\boxed{\text{customers have to ask}}$



X = NUMBER OF CUSTOMERS WHO LIKE VINEGAR

$X \sim B(20, 0.35)$

a) $H_0 : p = 0.35$

$H_1 : p < 0.35$

p IS THE PROPORTION OF ALL CUSTOMERS WHO LIKE VINEGAR

TESTING AT 5% SIGNIFICANCE, ON THE BASIS THAT $x = 3$

$P(X \leqslant 3) = 0.044375\ldots$

$= 4.4\%$

$< 5\%$

THERE IS SIGNIFICANT EVIDENCE TO SUPPORT EMMA'S CLAIM

THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

b) POSSIBLY BECAUSE THE VINEGAR IS NOT "FREELY AVAILABLE", I.E. CUSTOMERS HAVE TO ASK FOR VINEGAR

Question 5

The probability that a coffee vending machine will spill the drink is 25%.

The machine is now serviced, and after the service the next twenty dispenses of drinks are recorded.

It is found that only one drink is now spilled.

Test, at the 1% level of significance, whether the proportion of spilled drinks has reduced.

$\boxed{\text{not significant evidence, } 2.43\% > 1\%}$

$\boxed{\begin{array}{l} X = \text{NUMBER OF OVERFILLED CUPS} \\ X \sim B(20, 0.25) \end{array}}$

$H_0 : p = 0.25$

$H_1 : p < 0.25$

p IS THE PROPORTION OF ALL OVERFILLED CUPS BY THIS VENDING MACHINE

TESTING AT THE 1% SIGNIFICANCE LEVEL, ON THE BASIS THAT $x = 1$

$$P(X \leq 1) = 0.0243126\ldots = 0.024\ldots = 2.4\% > 1\%$$

THERE IS NO SIGNIFICANT EVIDENCE THAT THE PROPORTION OF SPILLED DRINKS HAS REDUCED

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 6

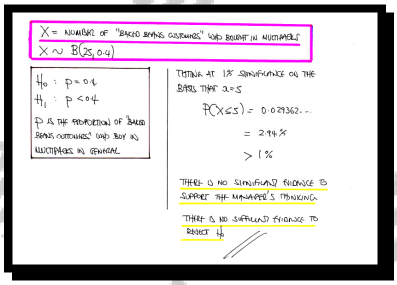
It was suggested to the new manager of a supermarket that 40% of the customers who buy baked beans will buy beans in "multi-packs".

The manager thinks that this figure is too high and decides to test it.

He finds that out of the next 25 customers who bought baked beans, only 5 customers bought baked beans in "multi-packs".

Test, at the 1% level of significance, whether or not there is evidence to support the supermarket's manager's thinking.

$\boxed{\text{not significant evidence, } 2.94\% > 1\%}$



X = NUMBER OF "BAKED BEANS CUSTOMERS" WHO BOUGHT IN MULTIPACKS

$X \sim B(25, 0.4)$

$H_0 : p = 0.4$

$H_1 : p < 0.4$

p IS THE PROPORTION OF "BAKED BEANS CUSTOMERS" WHO BUY IN MULTIPACKS IN GENERAL

TESTING AT 1% SIGNIFICANCE ON THE BASIS THAT $x = 5$

$P(X \leq 5) = 0.029362...$

$= 2.94\%$

$> 1\%$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE MANAGER'S THINKING

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 7

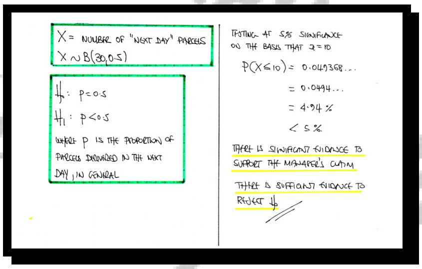
A parcel dispatch company has established that the probability of a parcel being delivered in the next working day is 0.5.

The company implements changes and the manager of a depot feels that fewer parcels are now delivered in the next working day.

He monitors a random sample of 30 parcels left to be delivered and finds 10 were delivered in the next working day.

Test, at the 5% level of significance, whether or not there is evidence to support the manager's claim.

$\boxed{\text{significant evidence, } 4.94\% < 5\%}$


$X =$ NUMBER OF "NEXT DAY" PARCELS

$X \sim B(30, 0.5)$

$H_0: p = 0.5$

$H_1: p < 0.5$

WHERE p IS THE PROPORTION OF PARCELS DELIVERED IN THE NEXT DAY, IN GENERAL

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 10$

$P(X \leq 10) = 0.049368\ldots$

$= 0.0494\ldots$

$= 4.94\%$

$< 5\%$

THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE MANAGER'S CLAIM

THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

Question 8

Johnny is often late for his lessons and according to his Head of Year this proportion is unacceptably high at 0.4 .

After a formal discussion for this matter, the Head of Year decides to check Johnny's punctuality by looking at a random sample of 30 lessons.

Johnny was late on 8 occasions in the sample.

Test, at the 5% level of significance, whether there is evidence that the discussion with the Head of Year has improved Johnny's punctuality.

$\boxed{\text{not significant evidence, } 9.4\% > 5\%}$

$\boxed{\begin{array}{l} X = \text{NUMBER OF "LATES"} \\ X \sim B(30, 0.4) \end{array}}$

$H_0 : p = 0.4$

$H_1 : p < 0.4$

WHERE p IS THE PROPORTION OF "LATES" IN GENERAL

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 8$

$P(X \leq 8) = 0.0940$

$= 9.4\%$

$> 5\%$

THERE IS NO SIGNIFICANT EVIDENCE THAT THE DISCUSSION HAS IMPROVED JOHNNY'S PUNCTUALITY

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 9

A pub manager feels that since the introduction of the "smoking ban" in his pub, the proportion of the non smoking customers visiting his pub has increased.

It had been established that before the "smoking ban" 15% of the customers visiting his pub were non smokers.

He now finds 8 non smokers in a random sample of 20 customers.

Test, at the 1% level of significance, whether there is evidence to suggest that the proportion of the non smoking customers visiting his pub has increased.

$\boxed{\text{significant evidence, } 0.59\% < 1\%}$

X = NUMBER OF NON SMOKING CUSTOMERS

$X \sim B(20, 0.15)$

- $H_0: p = 0.15$
- $H_1: p > 0.15$, WHERE p IS THE GENERAL PROPORTION OF NON SMOKERS VISITING THE PUB (POPULATION)

TESTING AT 1% SIGNIFICANCE ON THE BASIS THAT $x = 8$

$$
\begin{aligned}
P(X \geqslant 8) &= 1 - P(X \leqslant 7) \\
&= 1 - 0.9941 \quad \text{(tables)} \\
&= 0.0059 \\
&= 0.59\% < 1\%
\end{aligned}
$$

THERE IS SIGNIFICANT EVIDENCE TO SUGGEST THAT THE PROPORTION OF NON SMOKERS HAS INCREASED

THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

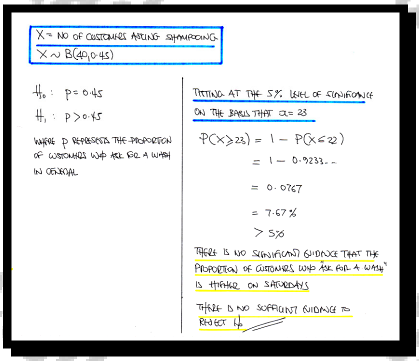
Question 10

In a salon 45% of the customers ask for the hair to be washed before it is cut.

During a busy Saturday 40 customers visited the salon and 23 of these customers asked for a wash.

Test, at the 5% level of significance, whether the proportion of customers who ask for their hair to be washed is higher on Saturdays.

$\boxed{\text{not significant evidence, } 7.67\% > 5\%}$



X = NO OF CUSTOMERS ASKING SHAMPOOING

$X \sim B(40, 0.45)$

$H_0 : p = 0.45$

$H_1 : p > 0.45$

WHERE p REPRESENTS THE PROPORTION OF CUSTOMERS WHO ASK FOR A WASH IN GENERAL

TESTING AT THE 5% LEVEL OF SIGNIFICANCE ON THE BASIS THAT $x = 23$

$P(X \geq 23) = 1 - P(X \leq 22)$
$= 1 - 0.9233...$
$= 0.0767$
$= 7.67\%$
$> 5\%$

THERE IS NO SIGNIFICANT EVIDENCE THAT THE PROPORTION OF CUSTOMERS WHO "ASK FOR A WASH" IS HIGHER ON SATURDAYS

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

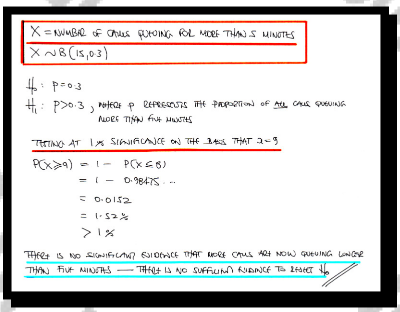
Question 11

In a certain bank, the probability that a phone call is in a queue for more than five minutes is 0.3.

A new telephone console is installed in order to improve efficiency, however when a sample of 15 calls is checked, 9 calls were found to be queuing for more than five minutes.

Test, at the 1% level of significance, whether there is evidence that more calls are now queuing for more than five minutes.

$\boxed{\text{not significant evidence, } 1.52\% > 1\%}$



X = NUMBER OF CALLS QUEUING FOR MORE THAN 5 MINUTES

$X \sim B(15, 0.3)$

$H_0 : p = 0.3$

$H_1 : p > 0.3$, WHERE p REPRESENTS THE PROPORTION OF ALL CALLS QUEUING MORE THAN FIVE MINUTES

TESTING AT 1% SIGNIFICANCE ON THE BASIS THAT $x = 9$

$P(X \geq 9) = 1 - P(X \leq 8)$
$= 1 - 0.98475...$
$= 0.0152$
$= 1.52\%$
$> 1\%$

THERE IS NO SIGNIFICANT EVIDENCE THAT MORE CALLS ARE NOW QUEUING LONGER THAN FIVE MINUTES — THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

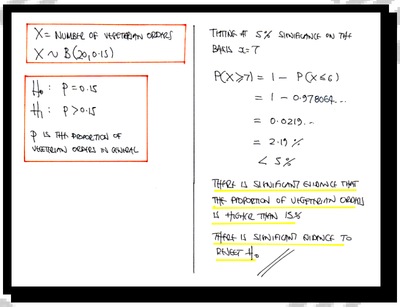
Question 12

The proportion of vegetarian orders in a restaurant is thought to be 15%.

During lunch on a given day, 7 diners out of 20, ordered vegetarian food.

Test, at 5% level of significance, whether the proportion of diners who order vegetarian food is higher than 15%.

$\boxed{\text{significant evidence, } 2.19\% < 5\%}$



X = NUMBER OF VEGETARIAN ORDERS

$X \sim B(20, 0.15)$

$H_0 : p = 0.15$

$H_1 : p > 0.15$

p IS THE PROPORTION OF VEGETARIAN ORDERS IN GENERAL

TESTING AT 5% SIGNIFICANCE ON THE BASIS $x = 7$

$P(X \geqslant 7) = 1 - P(X \leqslant 6)$
$= 1 - 0.978064...$
$= 0.0219...$
$= 2.19\%$
$< 5\%$

THERE IS SIGNIFICANT EVIDENCE THAT THE PROPORTION OF VEGETARIAN ORDERS IS HIGHER THAN 15%

THERE IS SIGNIFICANT EVIDENCE TO REJECT H_0

Question 13

A tomato plant is thought to produce a good yield if it produces more than 3 kg of fruit.

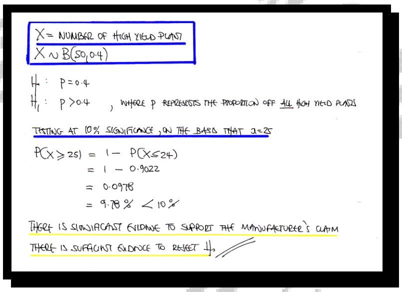
A farmer has established over a long period of time that the probability that a tomato plant will produce a good yield is 0.4 .

The makers of a new fertilizer claim that their product will increase the yield of the tomato plants.

The farmer uses the fertilizer in a random plot of 50 plants and finds half the plants in this plot produce a good yield.

Test, at the 10% level of significance, whether there is evidence to support the claim made by the makers of the fertilizer.

$\boxed{\text{significant evidence, } 9.78\% < 10\%}$



X = NUMBER OF HIGH YIELD PLANT
$X \sim B(50, 0.4)$

$H_0 : p = 0.4$
$H_1 : p > 0.4$, WHERE p REPRESENTS THE PROPORTION OFF ALL HIGH YIELD PLANTS

TESTING AT 10% SIGNIFICANCE, ON THE BASIS THAT $x = 25$

$$P(X \geqslant 25) = 1 - P(X \leqslant 24) = 1 - 0.9022 = 0.0978 = 9.78\% < 10\%$$

THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE MANUFACTURER'S CLAIM
THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

Question 14

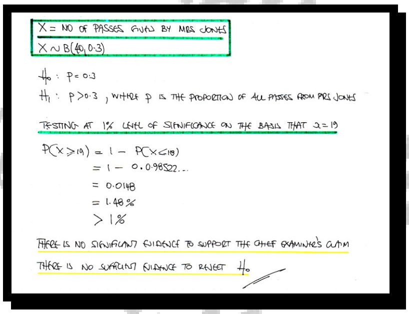
It had been established over a long period of time that 30% of the driving test candidates examined by Mrs Jones will pass their driving test.

The chief examiner of the centre, on his return after a year long break, feels that a higher proportion of candidates pass their driving test when examined by Mrs Jones.

In a random sample of 40, of Mrs Jones' recent candidates, he finds 19 passed their test.

Test, at the 1% level of significance, whether or not there is evidence to support the chief examiner's claim.

$\boxed{\text{not significant evidence, } 1.48\% > 1\%}$


X = NO OF PASSES GIVEN BY MRS JONES

$X \sim B(40, 0.3)$

$H_0 : p = 0.3$

$H_1 : p > 0.3$, WHERE p IS THE PROPORTION OF ALL PASSES FROM MRS JONES

TESTING AT 1% LEVEL OF SIGNIFICANCE ON THE BASIS THAT $x = 19$

$$\begin{aligned} P(X \geq 19) &= 1 - P(X \leq 18) \\ &= 1 - 0.0.98522\ldots \\ &= 0.0148 \\ &= 1.48\% \\ &> 1\% \end{aligned}$$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE CHIEF EXAMINER'S CLAIM

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 15

In a craft activity in a primary school, kids use beads which are kept in a bag. The bag contains a large number of beads of different colours. It is known that 0.3 of the beads are coloured gold.

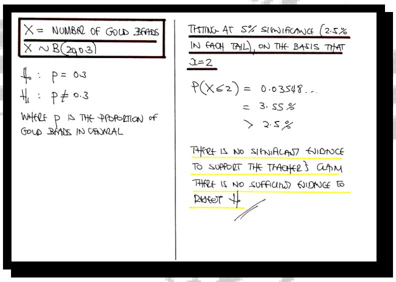
The teacher claims the proportion of gold beads in the bag has changed after the activity.

She checks a random sample of 20 beads out of the bag, after the end of the activity

She finds two gold beads in the sample.

Test, at the 5% level of significance, whether or not there is evidence to support the teacher's claim.

$\boxed{\text{not significant evidence, } 3.55\% > 2.5\%}$



X = NUMBER OF GOLD BEADS

$X \sim B(20, 0.3)$

$H_0 : p = 0.3$

$H_1 : p \neq 0.3$

WHERE p IS THE PROPORTION OF GOLD BEADS IN GENERAL

TESTING AT 5% SIGNIFICANCE (2.5% IN EACH TAIL), ON THE BASIS THAT $x=2$

$P(X \leqslant 2) = 0.03548\ldots$

$= 3.55\%$

$> 2.5\%$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE TEACHER'S CLAIM

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 16

It is claimed that 20% of the seeds sold by a nursery will germinate.

Ann buys 40 such seeds and finds that 12 of them germinated.

Test, at the 10% level of significance, whether or not there is evidence to support the nursery's claim.

$\boxed{\text{not significant evidence, } 8.75\% > 5\%}$

$\boxed{\begin{array}{l} X = \text{NUMBER OF SEEDS WHICH GERMINATE} \\ X \sim B(40, 0.2) \end{array}}$

$H_0 : p = 0.2$

$H_1 : p \neq 0.2$, WHERE p IS THE PROPORTION OF GERMINATING SEEDS IN GENERAL (I.E. IN THE POPULATION)

TESTING AT 10% SIGNIFICANCE, TWO TAILED, SO 5% IN EACH TAIL, ON THE BASIS THAT $x = 12$

$$\begin{aligned} P(X \geqslant 12) &= 1 - P(X \leqslant 11) \\ &= 1 - 0.9125 \\ &= 0.0875 \\ &= 8.75\% \\ &> 5\% \end{aligned}$$

THERE IS NO SIGNIFICANT EVIDENCE, IF THERE IS EVIDENCE TO SUPPORT THE NURSERY'S CLAIM

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

BINOMIAL DISTRIBUTION HYPOTHESIS

EXAM QUESTIONS

Question 1 (***)

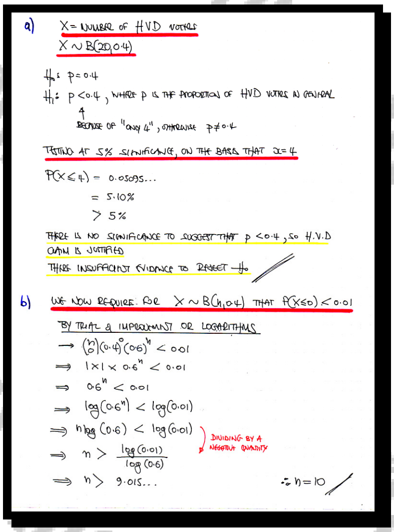
A mayoral candidate, Hans Van Dyke, claims that 40% of the electoral will vote for him in the next election. In a recent opinion poll of 20 recently selected voters it was found that **only 4** people will vote for Hans Van Dyke.

a) Test, at the 5% level of significance whether, or not, the opinion poll supports Hans Van Dyke's claim.

In a second opinion poll of n randomly selected people, it was found that no one will be voting for Hans Van Dyke. As a result of this poll, Hans Van Dyke's claim is rejected at 1% significance.

b) Determine the smallest value of n.

$\boxed{}$, $\boxed{\text{not significant evidence, } 5.10\% > 5\%}$, $\boxed{n = 10}$



a) X = NUMBER OF HVD VOTERS

$X \sim B(20, 0.4)$

$H_0: p = 0.4$

$H_1: p < 0.4$, WHERE p IS THE PROPORTION OF HVD VOTERS IN GENERAL

BECAUSE OF "ONLY 4", OTHERWISE $p \neq 0.4$

TESTING AT 5% SIGNIFICANCE, ON THE BASIS THAT $x = 4$

$P(X \leqslant 4) = 0.05095\ldots$

$= 5.10\%$

$> 5\%$

THERE IS NO SIGNIFICANCE TO SUGGEST THAT $p < 0.4$, SO H.V.D CLAIM IS JUSTIFIED

THERE INSUFFICIENT EVIDENCE TO REJECT H_0

b) WE NOW REQUIRE FOR $X \sim B(n, 0.4)$ THAT $P(X \leqslant 0) < 0.01$

BY TRIAL & IMPROVEMENT OR LOGARITHMS

$\rightarrow \binom{n}{0}(0.4)^0(0.6)^n < 0.01$

$\Rightarrow 1 \times 1 \times 0.6^n < 0.01$

$\Rightarrow 0.6^n < 0.01$

$\Rightarrow \log(0.6^n) < \log(0.01)$

$\Rightarrow n\log(0.6) < \log(0.01)$

DIVIDING BY A NEGATIVE QUANTITY

$\Rightarrow n > \dfrac{\log(0.01)}{\log(0.6)}$

$\Rightarrow n > 9.015\ldots$

$\therefore n = 10$

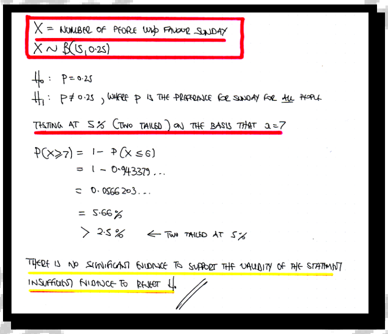
Question 2 (***)

When people are asked "what is your favourite day of the week?", it is thought that on average one person in four replies "Sunday".

To test this assertion 15 people were asked this question and 7 replied "Sunday".

Carry out a significance test, at the 5% level, of whether the statement "on average the preferred day of the week is Sunday, for one in four persons".

$\boxed{}$, $\boxed{\text{not significant evidence, } 5.66\% > 2.5\%}$



$\boxed{\begin{array}{l} X = \text{NUMBER OF PEOPLE WHO FAVOUR SUNDAY} \\ X \sim B(15, 0.25) \end{array}}$

$H_0 : p = 0.25$

$H_1 : p \neq 0.25$, WHERE p IS THE PREFERENCE FOR SUNDAY FOR ALL PEOPLE

TESTING AT 5% (TWO TAILED) ON THE BASIS THAT $x = 7$

$$\begin{aligned} P(X \geqslant 7) &= 1 - P(X \leqslant 6) \\ &= 1 - 0.943379\ldots \\ &= 0.0566203\ldots \\ &= 5.66\% \\ &> 2.5\% \quad \leftarrow \text{TWO TAILED AT } 5\% \end{aligned}$$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE VALIDITY OF THE STATEMENT

INSUFFICIENT EVIDENCE TO REJECT H_0

Question 3 (***)

A chicken farmer supplies a local restaurant with eggs. The owner of the restaurant feels that the eggs having a double yolk is 0.009.

The farmer however claims that the proportion of her eggs having a double yolk is higher than 0.009.

In the next batch of 500 eggs the restaurant chef records 9 eggs with a double yolk.

Using a distributional approximation, and further assuming that the batch of 500 eggs is a random sample, test at the 5% level of significance whether the farmer's claim is justified.

$\boxed{}$, $\boxed{\text{claim justified, } 4.03\% < 5\%}$

X = NUMBER OF EGGS WITH DOUBLE YOLK

$X \sim B(500, 0.009)$

H_0 $p = 0.009$

H_1 : $p > 0.009$, WHERE p IS THE PROPORTION OF EGGS WITH DOUBLE YOLK IN GENERAL

TESTING AT 5% SIGNIFICANCE, ON THE BASIS THAT $x = 9$

$P(X \geqslant 9) = 1 - P(X \leqslant 8)$

USING POISSON TABLES

$= 1 - 0.95974\ldots$

$= 0.04025\ldots$

$= 4.03\%$

$< 5\%$

AS n IS LARGE & p IS SMALL APPROXIMATE BY

$X \sim Po(np)$

IE $X \sim Po(4.5)$

500×0.009

THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE FARMER'S CLAIM

SUFFICIENT EVIDENCE TO REJECT H_0

Question 4 (***+)

A London Mayoral candidate thinks that 35% of the voters will vote for him.

His campaign manager thinks that this percentage is in fact higher.

A random sample of 40 voters is polled and it is found that it contained 19 voters indicating their intention to vote in favour of this candidate.

a) Test at the 5% significance level the claim of the campaign manager.

A larger random sample of 200 is then polled. It is found that it contained 84 voters indicating their intention to vote in favour of this candidate.

b) Use a distributional approximation, to retest at the 5% significance level the claim of the campaign manager.

$\boxed{}$, $\boxed{\text{not significant evidence, } 6.99\% > 5\%}$, $\boxed{\text{significant evidence, } 2.27\% < 5\%}$

a) X = NUMBER OF VOTERS IN FAVOUR

$X \sim B(40, 0.35)$

$H_0: p = 0.35$

$H_1: p > 0.35$, WHERE p IS THE PROPORTION OF "IN FAVOUR" VOTERS IN GENERAL

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 19$

$$P(X \geqslant 19) = 1 - P(X \leqslant 18) = 1 - 0.93008\ldots = 0.0699 = 6.99\% > 5\%$$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE CAMPAIGN MANAGER'S CLAIM — NO SUFFICIENT EVIDENCE TO REJECT H_0

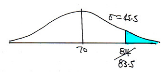
b) NOW SAMPLE $n = 200$, HYPOTHESES EXACTLY THE SAME

$X \sim B(200, 0.35)$

$P(X \geqslant 84) = \ldots$ NEED NOW TO BE FOUND BY APPROXIMATION

$Y \sim N(np, np(1-p))$

$Y \sim N(70, 45.5)$



$$= P(Y > 83.5) = 1 - P(Y < 83.5) = 1 - P\left(Z < \frac{83.5 - 70}{\sqrt{45.5}}\right) = 1 - \Phi(2.00137\ldots) = 1 - 0.97732\ldots = 0.02267 = 2.27\% < 5\%$$

THERE IS NOW SIGNIFICANT EVIDENCE TO SUPPORT THE CAMPAIGN MANAGER'S CLAIM — THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

Question 5 (***+)

Drawing pins are sold in boxes of 20 and it is thought that 10% of these drawing pins have flaws.

a) If a box of these drawing pins is examined at random, find the probability that it will contain ...

i. ... 3 drawing pins with flaws.

ii. ... at least 2 drawing pins with flaws.

b) If 3 boxes of these drawing pins are examined at random, determine the probability that at least one of these boxes will contain at least 2 drawing pins with flaws.

A single box of these drawing pins is picked at random and found to contain exactly 5 drawing pins with flaws.

c) Test, at the 5% level of significance, whether this constitutes evidence that the proportion of drawing pins with flaws is higher than 10% .

$\boxed{}$, $\boxed{0.1901}$, $\boxed{0.6083}$, $\boxed{0.9399}$, $\boxed{\text{significant, } 4.32\% < 5\%}$

a) $\boxed{\begin{array}{l} X = \text{NUMBER OF PINS WITH FLAWS} \\ X \sim B(20, 0.1) \end{array}}$

I) $P(X=3) = \binom{20}{3}(0.1)^3(0.9)^{17} \simeq 0.1901$

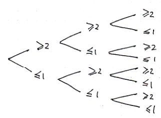
II) $P(X \geq 2) = 1 - P(X \leq 1)$ ... tables

$= 1 - 0.3917$

$= 0.6083$

b) USING ALL THE OUTCOMES OR BY SETTING ANOTHER DISTRIBUTION OR A TREE DIAGRAM

$\boxed{P(X \geq 2) = 0.6083}$ & $\boxed{P(X \leq 1) = 0.3917}$



THE REQUIRED PROBABILITY IS GIVEN BY ALL THE BRANCHES ABOVE EXCEPT THE BOTTOM ONE

OR $1 - [P(X \leq 1)]^3 = 1 - 0.3917^3 = 0.9399$

c) SETTING UP HYPOTHESES

- $H_0 : p = 0.1$
- $H_1 : p > 0.1$
- WHERE p REPRESENTS THE PROPORTION OF PINS WITH FLAWS IN THE ENTIRE POPULATION & NOT IN THE SAMPLE

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $X=5$

$P(X \geq 5) = 1 - P(X \leq 4)$

... tables...

$= 1 - 0.9568$

$= 0.0432$

$= 4.32\ \%$

$< 5\%$

THERE IS SIGNIFICANT EVIDENCE AT 5% THAT THE PROPORTION OF PINS WITH FLAWS IS HIGHER THAN 10%

WE DO NOT HAVE ENOUGH EVIDENCE TO REJECT H_1

Question 6 (***+)

The discrete random variable X represents the number of households with satellite TV subscriptions.

It is assumed that X follows a binomial distribution $\mathrm{B}(n, 0.35)$.

a) If $n = 25$, find the probability ...

i. ... $\mathrm{P}(X = 12)$

ii. ... $\mathrm{P}(X > 12)$.

b) If $n = 25$, determine the probability

$$\mathrm{P}\left[\mathrm{E}(X) - \sqrt{\mathrm{Var}(X)} < X < \mathrm{E}(X) + \sqrt{\mathrm{Var}(X)}\right].$$

c) Find the smallest number of households that must be sampled so that the probability of having at least a household with satellite TV subscription is greater than 99%.

An analyst believes that the proportion of households with satellite TV subscription is higher, because in a sample of 25 households 13 had a satellite TV subscription.

d) Using a clear method, test the analyst's belief, at the 5% level of significance.

$\boxed{}$, $\boxed{0.0650}$, $\boxed{0.0604}$, $\boxed{0.7012}$, $\boxed{n = 12}$, $\boxed{\text{not significant, } 6.04\% > 5\%}$

X = NUMBER OF "PEOPLE" WITH SATELLITE SUBSCRIPTIONS
$X \sim B(25, 0.35)$

a)I) $P(X=12) = \binom{25}{12}(0.35)^{12}(0.65)^{13} = 0.064971.. \approx 0.0650$

II) $P(X>12) = P(X \geq 13) = 1 - P(X \leq 12)$

tables or calculator

$= 1 - 0.9396$
$= 0.0604$

b) START BY FINDING THE EXPECTATION & VARIANCE

$E(X) = np = 25 \times 0.35 = 8.75$

$Var(X) = np(1-p) = 8.75 \times 0.65 = 5.6875$

HENCE WE OBTAIN

$P\left[E(X) - \sqrt{Var(X)} < X < E(X) + \sqrt{Var(X)}\right]$
$= P[8.75 - 2.3848... < X < 8.75 + 2.3848...]$
$= P(6.365... < X < 11.135..)$
$= P(7 \leq X \leq 11)$
$= P(X \leq 11) - P(X \leq 6)$

tables (or calculator)

$= 0.8746 - 0.1734$
$= 0.7012$

c) LET THE REQUIRED SAMPLE BE N

$\Rightarrow P(X \geq 1) > 99\%$
$\Rightarrow P(X = 0) < 1\%$
$\Rightarrow \binom{N}{0}(0.35)^0(0.65)^N < 0.01$
$\Rightarrow 1 \times 1 \times 0.65^N < 0.01$

BY LOGS (OR TRIAL & IMPROVEMENT)

$\Rightarrow 0.65^N < 0.01$
$\Rightarrow \log(0.65^N) < \log(0.01)$
$\Rightarrow N\log(0.65) < \log(0.01)$
$\Rightarrow N > \dfrac{\log(0.01)}{\log(0.65)}$

$\log(0.65)$ IS NEGATIVE SO THE INEQUALITY REVERSES

$\Rightarrow N > 10.6902..$

$\therefore N = 11$

d) SETTING UP HYPOTHESES

- $H_0: p = 0.35$
- $H_1: p > 0.35$, WHERE p REPRESENTS THE PROPORTION OF HOUSEHOLDS WITH SATELLITE TV IN THE POPULATION (NOT THE SAMPLE)

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 13$

$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.9396 = 0.0604$
$= 6.04\% > 5\%$

THERE IS NO SIGNIFICANT EVIDENCE THAT THE PROPORTION OF HOUSEHOLDS WITH SATELLITE T.V IS HIGHER THAN 35%

THERE IS NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 7 **(****)**

William has been established over a long period of time that when he shoots an arrow at a target the probability of hitting it, is 0.35.

William buys a professional bow as he believes it will increase the probability of hitting the target.

Let X define the number of successful hits of the target, and assume that successful or non successful hits are independent of one another.

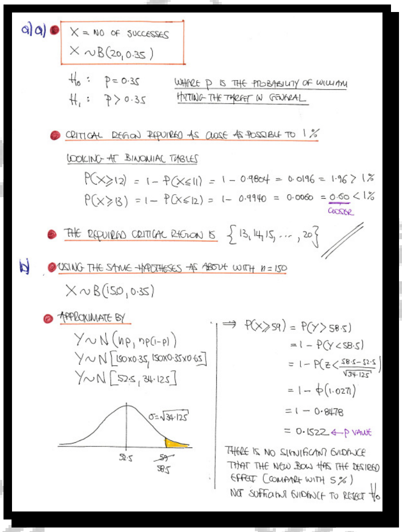
a) Determine a critical region based on 20 shots with the new bow.

The significance level must be as close as possible to 1%.

William decides to carry out a significance test by shooting 150 arrows at the target using the new bow. He finds that 59 arrows hit the target.

b) Using a distributional approximation, calculate an approximate p-value and hence state the conclusion in context.

$\boxed{}$, $\boxed{\{13,14,15,...,20\}}$, $\boxed{p = 0.1522}$, $\boxed{\text{not significant evidence, } 15.22\% > 5\%}$


a) a) $\boxed{\begin{array}{l} X = \text{NO OF SUCCESSES} \\ X \sim B(20, 0.35) \end{array}}$

$H_0: p = 0.35$
$H_1: p > 0.35$
WHERE p IS THE PROBABILITY OF WILLIAM HITTING THE TARGET IN GENERAL

CRITICAL REGION REQUIRED AS CLOSE AS POSSIBLE TO 1%

LOOKING AT BINOMIAL TABLES

$P(X \geqslant 12) = 1 - P(X \leqslant 11) = 1 - 0.9804 = 0.0196 = 1.96 > 1\%$
$P(X \geqslant 13) = 1 - P(X \leqslant 12) = 1 - 0.9940 = 0.0060 = 0.60 < 1\%$
CLOSER

THE REQUIRED CRITICAL REGION IS $\{13, 14, 15, \ldots, 20\}$

b) USING THE SAME HYPOTHESES AS ABOVE WITH $n = 150$

$X \sim B(150, 0.35)$

APPROXIMATE BY

$Y \sim N(np, np(1-p))$
$Y \sim N[150 \times 0.35, 150 \times 0.35 \times 0.65]$
$Y \sim N[52.5, 34.125]$

$\sigma = \sqrt{34.125}$
52.5
~~59~~ 58.5

$\Rightarrow P(X \geqslant 59) = P(Y > 58.5)$
$= 1 - P(Y < 58.5)$
$= 1 - P\left(Z < \frac{58.5 - 52.5}{\sqrt{34.125}}\right)$
$= 1 - \Phi(1.0271)$
$= 1 - 0.8478$
$= 0.1522 \leftarrow p \text{ VALUE}$

THERE IS NO SIGNIFICANT EVIDENCE THAT THE NEW BOW HAS THE DESIRED EFFECT (COMPARE WITH 5%)
NOT SUFFICIENT EVIDENCE TO REJECT H_0

Question 8 (****)

At "Stavros Restaurant" the owner is told by his chef that 25% of the customers order vegetarian food. The owner wants to check the validity of the chef's assertion so he checks a random sample of 20 orders, only to find 2 vegetarian orders.

a) Is there evidence, at the 10% level of significance, that the proportion of vegetarian orders is lower than 25%?

At "Mavros Restaurant" the owner is told by his waiters that 25% of the customers order vegetarian food. The owner wants to check the validity of the waiters' belief so he checks a random sample of 100 orders.

b) Given that there are 31 vegetarian orders in the sample, use a distributional approximation, to test at the 5% level of significance, the belief of the waiters at "Mavros Restaurant".

$\boxed{\quad}$, $\boxed{\text{significant, } 9.13\% < 10\%}$, $\boxed{\text{"not significant!!!", } 10.20\% > 2.5\%}$

a) $\boxed{\begin{array}{l} X = \text{NUMBER OF VEGETARIAN ORDERS} \\ X \sim B(20, 0.25) \end{array}}$

$H_0: p = 0.25$

$H_1: p < 0.25$, WHERE p IS THE PROPORTION OF VEGETARIAN ORDERS IN GENERAL

TESTING AT 10% SIGNIFICANCE ON THE BASIS THAT $x = 2$

$$P(X \leq 2) = 0.09126\ldots = 9.13\% < 10\%$$

THERE IS SIGNIFICANT EVIDENCE THAT THE PROPORTION OF VEGETARIAN ORDERS IS LOWER THAN 25%

THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

b) NOW SAMPLE IS 100

$H_0: p = 0.25$

$H_1: p \neq 0.25$, WHERE p IS THE PROPORTION OF VEGETARIAN ORDERS IN GENERAL

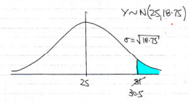
TESTING AT 5% SIGNIFICANCE, ON THE BASIS $x = 31$ (TWO TAILED)

APPROXIMATE BY NORMAL

$E(X) = \text{MEAN} = np = 100 \times 0.25 = 25$

$\text{Var}(X) = \text{VARIANCE} = np(1-p) = 25 \times 0.75 = 18.75$

$Y \sim N(25, 18.75)$



$$\begin{aligned} P(X \geq 31) &= P(Y > 30.5) \\ &= 1 - P(Y < 30.5) \\ &= 1 - P\left(Z < \frac{30.5 - 25}{\sqrt{18.75}}\right) \\ &= 1 - \Phi(1.27017\ldots) \\ &= 1 - 0.897988\ldots \\ &= 0.1020\ldots \\ &= 10.2\% \\ &> 2.5\% \end{aligned}$$

THERE IS NON SIGNIFICANT EVIDENCE TO SUPPORT THE WAITERS' CLAIM

INSUFFICIENT EVIDENCE TO REJECT H_0

BINOMIAL DISTRIBUTION CRITICAL REGIONS

Question 1 (**)

A nutritional expert believes that children ate healthier in the past. This is measured by the number of kids that eat five portions of fruit and vegetables per day. The proportion of children that ate "five a day" was 40% in the 1950's.

The nutritional expert believes that the proportion is now lower.

He carries out an investigation with a random sample of 15 children.

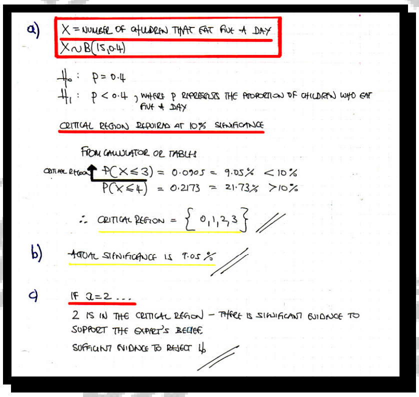
a) Find the critical region to test at the 10% level of significance the expert's belief.

b) State the actual significance level for a test using the critical region of part **(a)**.

Two children, who ate "five a day", were found in the sample.

c) Complete the test.

$\boxed{\text{C.R.} = \{0,1,2,3\}}$, $\boxed{9.05\%}$, $\boxed{\text{significant evidence, } 2 \text{ is in C.R.}}$



a) $\boxed{X = \text{NUMBER OF CHILDREN THAT EAT FIVE A DAY} \\ X \sim B(15, 0.4)}$

$H_0: p = 0.4$

$H_1: p < 0.4$, WHERE p REPRESENTS THE PROPORTION OF CHILDREN WHO EAT FIVE A DAY

CRITICAL REGION REQUIRED AT 10% SIGNIFICANCE

FROM CALCULATOR OR TABLES

CRITICAL REGION $P(X \leq 3) = 0.0905 = 9.05\% < 10\%$

$P(X \leq 4) = 0.2173 = 21.73\% > 10\%$

$\therefore$ CRITICAL REGION $= \{0,1,2,3\}$ //

b) ACTUAL SIGNIFICANCE IS 9.05% //

c) IF $x = 2$ ...

2 IS IN THE CRITICAL REGION – THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE EXPERT'S BELIEF

SUFFICIENT EVIDENCE TO REJECT H_0 //

Question 2 (**)

The proportion of tiles with minor faults produced in a factory is thought to be 10%.

The factory manager believes that the proportion is higher due to the old machinery.

He inspects a random sample of 20 tiles.

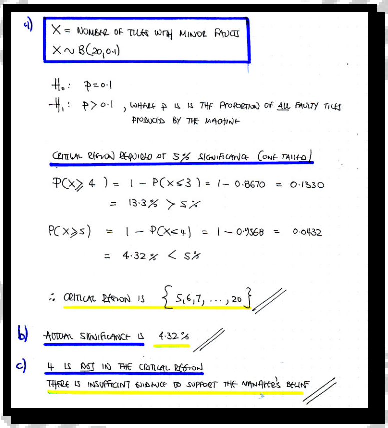
a) Find the critical region to test the manager's belief, at the 5% level of significance.

b) State the actual significance level for a test using the critical region of part **(a)**.

Four faulty tiles were found in the sample.

c) Complete the test.

$\boxed{\quad}$, $\boxed{\text{C.R.} = \{5,6,7,\ldots,20\}}$, $\boxed{4.32\%}$, $\boxed{\text{not significant evidence, } 4 \text{ not in C.R.}}$



a) $\boxed{\begin{array}{l} X = \text{NUMBER OF TILES WITH MINOR FAULTS} \\ X \sim B(20, 0.1) \end{array}}$

$H_0: p = 0.1$

$H_1: p > 0.1$, WHERE p IS IS THE PROPORTION OF ALL FAULTY TILES PRODUCED BY THE MACHINE

CRITICAL REGION REQUIRED AT 5% SIGNIFICANCE (ONE TAILED)

$P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.8670 = 0.1330$
$= 13.3\% > 5\%$

$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.9568 = 0.0432$
$= 4.32\% < 5\%$

$\therefore$ CRITICAL REGION IS $\{5, 6, 7, \ldots, 20\}$ //

b) ACTUAL SIGNIFICANCE IS 4.32% //

c) 4 IS NOT IN THE CRITICAL REGION

THERE IS INSUFFICIENT EVIDENCE TO SUPPORT THE MANAGER'S BELIEF //

Question 3 (**)

The recruitment director of a large accounting firm believes that maths graduates are more successful when applying for positions in his firm compared with graduates of other subjects.

One in five job applicants to this accounting firm is successful.

The recruitment director selects a random sample of 25 maths graduate applicants.

a) Find the critical region to test at the 5% level of significance the director's belief.

State your hypotheses clearly.

b) State the probability of incorrectly rejecting the null hypothesis in a test, using the critical region obtained in part **(a)**.

Ten successful maths graduate applicants were found in the sample.

c) Complete the test.

$\boxed{}$, $\boxed{\text{C.R.} = \{9, 10, 11, \ldots, 25\}}$, $\boxed{4.68\%}$, $\boxed{\text{significant evidence, } 10 \text{ is in C.R.}}$

a) X = NUMBER OF SUCCESSFUL APPLICANTS AMONGST MATHS GRADUATES

$X \sim B(25, 0.2)$

$H_0 : p = 0.2$

$H_1 : p > 0.2$, WHERE p IS THE PROPORTION OF SUCCESSFUL APPLICANTS IN GENERAL

CRITICAL REGION REQUIRED AT 5% SIGNIFICANCE

$P(X \geqslant 8) = 1 - P(X \leqslant 7) = 1 - 0.8908 = 0.1092 = 10.92\% > 5\%$

$P(X \geqslant 9) = 1 - P(X \leqslant \ ) = 1 - 0.9532 = 0.0468 = 4.68\% < 5\%$

$\therefore$ CRITICAL REGION $= \{9, 10, 11, \ldots, 25\}$

b) ACTUAL SIGNIFICANCE IS 4.68%

c) TEN IS IN THE CRITICAL REGION

THERE IS SUFFICIENT EVIDENCE TO SUPPORT THE DIRECTOR'S BELIEF

Question 4 (**)

Ten years ago the residents in a car congested area were asked whether they were in favour of a residents' parking scheme. The proportion of residents who supported the parking permit scheme was 30%. The scheme was never implemented due to lack of funding.

The funding is now available and a new councillor believes that the support for the scheme is different now.

The replies to questionnaires of twenty current residents are considered.

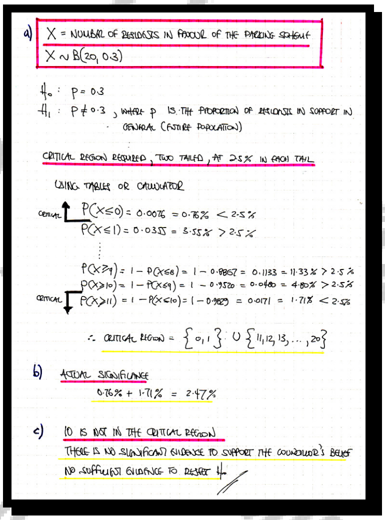
a) Determine the critical region to test, at the 5% level of significance, the councillor's belief.

b) State the actual significance level for a test using the critical region of part **(a)**.

Ten residents in support of the scheme were found in the sample.

c) Complete the test.

$\boxed{\text{C.R.} = \{0,1\} \cup \{11,12,13,...,20\}}$, $\boxed{2.47\%}$, $\boxed{\text{not significant evidence, 10 is not in C.R.}}$



a) X = NUMBER OF RESIDENTS IN FAVOUR OF THE PARKING SCHEME

$X \sim B(20, 0.3)$

$H_0 : p = 0.3$

$H_1 : p \neq 0.3$, WHERE p IS THE PROPORTION OF RESIDENTS IN SUPPORT IN GENERAL (ENTIRE POPULATION)

CRITICAL REGION REQUIRED, TWO TAILED, AT 2.5% IN EACH TAIL

USING TABLES OR CALCULATOR

CRITICAL $P(X \leq 0) = 0.0076 = 0.76\% < 2.5\%$

$P(X \leq 1) = 0.0355 = 3.55\% > 2.5\%$

$\vdots$

$P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.8867 = 0.1133 = 11.33\% > 2.5\%$

$P(X \geq 10) = 1 - P(X \leq 9) = 1 - 0.9520 = 0.0480 = 4.80\% > 2.5\%$

CRITICAL $P(X \geq 11) = 1 - P(X \leq 10) = 1 - 0.9829 = 0.0171 = 1.71\% < 2.5\%$

$\therefore$ CRITICAL REGION $= \{0, 1\} \cup \{11, 12, 13, \ldots, 20\}$

b) ACTUAL SIGNIFICANCE

$0.76\% + 1.71\% = 2.47\%$

c) 10 IS NOT IN THE CRITICAL REGION

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE COUNCILLOR'S BELIEF

NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 5 (**)

In a small factory, the quality of ceramic tiles is monitored daily by checking a random sample of 50 tiles from the day's production. It is required to find whether the proportion of tiles with minor faults is different from 10%.

a) Find the critical region for this test.

The probability of rejecting at either tail must be as close as possible to 2.5%.

b) State the actual significance level for a test using the critical region of part (**a**).

c) If nine faulty tiles are found in the sample complete the test.

$\boxed{\text{C.R.} = \{0,1\} \cup \{10,11,12,...,50\}}$, $\boxed{5.83\%}$, $\boxed{\text{not significant evidence, } 9 \text{ is not in C.R.}}$

a) X = NUMBER OF TILES WITH MINOR FAULTS

$X \sim B(50, 0.1)$

$H_0 : p = 0.1$

$H_1 : p \neq 0.1$, WHERE p IS THE PROPORTION OF FAULTY TILES IN GENERAL

CRITICAL REGION REQUIRED; REJECTION AS CLOSE AS POSSIBLE TO 2.5% IN EACH TAIL — CALCULATOR OR TABLES

CRITICAL REGION

$P(X \leqslant 0) = 0.0052 = 0.52\%$

$P(X \leqslant 1) = 0.0338 = 3.38\%$ ← CLOSER TO 2.5%

$\vdots$

CRITICAL REGION

$P(X \geqslant 9) = 1 - P(X \leqslant 8) = 1 - 0.9421 = 0.0579 = 5.79\%$

$P(X \geqslant 10) = 1 - P(X \leqslant 9) = 1 - 0.9755 = 0.0245 = 2.45\%$ ← CLOSER TO 2.5%

CRITICAL REGION $= \{0,1\} \cup \{10,11,12,\ldots,50\}$

b) ACTUAL SIGNIFICANCE $= 3.38\% + 2.45\% = 5.83\%$

c) NINE IS NOT IN THE CRITICAL REGION

THERE IS NO SIGNIFICANT EVIDENCE THAT THE PROPORTION OF FAULTY TILES IS DIFFERENT (HIGHER) THAN 10%

NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 6 (**)

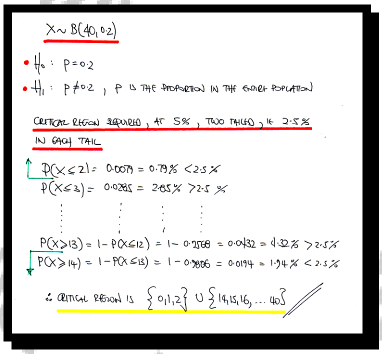
A test statistic has distribution $\mathrm{B}(40, p)$.

Given that

$$\mathrm{H}_0 : p = 0.2 , \quad \mathrm{H}_1 : p \neq 0.2 ,$$

find the critical region for the test statistic at the 5% significance level.

$\boxed{}$, $\boxed{\{0,1,2\} \cup \{14,15,16,\ldots,40\}}$



$X \sim B(40, 0.2)$

- $H_0 : p = 0.2$
- $H_1 : p \neq 0.2$, p IS THE PROPORTION IN THE ENTIRE POPULATION

CRITICAL REGION REQUIRED, AT 5%, TWO TAILED, IE 2.5% IN EACH TAIL

$P(X \leq 2) = 0.0079 = 0.79\% < 2.5\%$

$P(X \leq 3) = 0.0285 = 2.85\% > 2.5\%$

$\vdots$

$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.9568 = 0.0432 = 4.32\% > 2.5\%$

$P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.9806 = 0.0194 = 1.94\% < 2.5\%$

$\therefore$ CRITICAL REGION IS $\{0,1,2\} \cup \{14,15,16,\ldots 40\}$

Question 7 (**)

A test statistic has distribution $\mathrm{B}(30, p)$.

Given that

$$\mathrm{H}_0 : p = 0.4 , \quad \mathrm{H}_1 : p \neq 0.4 ,$$

find the critical region for the test statistic at the 5% significance level.

$\boxed{}$, $\boxed{\{0,1,2,3,4,5,6\} \cup \{18,19,20,\ldots,30\}}$

$X \sim B(30, 0.4)$

$H_0 : p = 0.4$

$H_1 : p \neq 0.4$, p PROPORTION OF THE POPULATION

CRITICAL REGION REQUIRED AT 5% SIGNIFICANCE, TWO TAILED SO 2.5% IN EACH TAIL

$P(X \leqslant 6) = 0.0172 = 1.72\% < 2.5\%$

$P(X \leqslant 7) = 0.0435 = 4.35\% > 2.5\%$

$\vdots$

$P(X \geqslant 17) = 1 - P(X \leqslant 16) = 1 - 0.9519 = 0.0481 = 4.81\% > 2.5\%$

$P(X \geqslant 18) = 1 - P(X \leqslant 17) = 1 - 0.9788 = 0.0212 = 2.12\% < 2.5\%$

$\therefore$ CRITICAL REGION IS

$\{0,1,2,\ldots,6\} \cup \{18,19,20,\ldots,30\}$

Question 8 (**+)

A test statistic has distribution $B(25, p)$.

a) Given that

$$H_0 : p = 0.35, \quad H_1 : p \neq 0.35,$$

find the critical region for the test statistic such that the probability of rejecting in each tail is as close as possible to 2.5%.

b) State the probability of incorrectly rejecting H_0 using this critical region.

$\boxed{}$, $\boxed{\{0,1,2,3,4\} \cup \{14,15,16,\ldots,25\}}$, $\boxed{5.75\% \text{ or } 0.0575}$

a) $X \sim B(25, 0.35)$

$H_0 : p = 0.35$

$H_1 : p \neq 0.35$, WHERE p IS THE PROPORTION IN GENERAL

CRITICAL REGION REQUIRED, TWO TAILED, REJECTING IN EACH TAIL AS CLOSE AS POSSIBLE TO 2.5%

- $P(X \leqslant 3) = 0.0097 = 0.97\%$
- $P(X \leqslant 4) = 0.0320 = 3.20\%$ ← CLOSER TO 2.5%

$\vdots$

NOT ACTUALLY NEEDED

- $P(X \geqslant 13) = 1 - P(X \leqslant 12) = 1 - 0.9396 = 0.0604 = 6.04\%$
- $P(X \geqslant 14) = 1 - P(X \leqslant 13) = 1 - 0.9745 = 0.0255 = 2.55\%$

CLOSER TO 2.5%

- $P(X \geqslant 15) = 1 - P(X \leqslant 14) = 1 - 0.9907 = 0.0093 = 0.93\%$

CRITICAL REGION IS $\{0,1,2,3,4\} \cup \{14,15,16,\ldots,25\}$

b) "ACTUAL SIGNIFICANCE" $= 3.20\% + 2.55\% = 5.75\%$

BINOMIAL DISTRIBUTION CRITICAL REGIONS

EXAM QUESTIONS

Question 1 (***+)

The records in a dentist's surgery show that 15% of the patients that make an appointment fail to turn up.

a) In a day with 15 appointments determine the probability that …

 i. … exactly 2 patients will fail to turn up.

 ii. … at least 2 patients will fail to turn up.

b) Use a distributional approximation to find the probability that in a month with 400 appointments, more than 45 but at most 65 patients will fail to turn up.

The surgery administrator feels that the percentage of patients that make an appointment and fail to turn up is likely to change in the future. The surgery tries an automated system of generating message reminders on patients' mobile phones.

It is required to find whether the proportion of patients that fail to turn up is different from 15%, by monitoring the next 20 appointments.

c) Determine the critical region for this test, at the 10% level of significance.

d) Write down the critical region for the same test if the probability of rejecting at either tail is as close as possible to 5%.

$\boxed{}$, $\boxed{0.2856}$, $\boxed{0.6814}$, $\boxed{0.7582}$, $\boxed{\{0\}\cup\{7,8,9,...,20\}}$, $\boxed{\{0\}\cup\{6,7,8,9,...,20\}}$

X = No of "No Show"

$X \sim B(15, 0.15)$

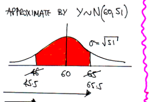
a) I) $P(X=2) = \binom{15}{2}(0.15)^2(0.85)^{13} = 0.2856$

II) $P(X \geqslant 2) = 1 - P(X \leqslant 1) = 1 - 0.3186 = 0.6814$

b) $X \sim B(400, 0.15)$

- $E(X) = np = 400 \times 0.15 = 60$
- $Var(X) = np(1-p) = 60 \times 0.85 = 51 > 5$

APPROXIMATE BY $Y \sim N(60, 51)$



$P(45 < X \leqslant 65)$

$= P(46 \leqslant X \leqslant 65)$

$= P(45.5 < Y < 65.5)$

$= P(Y < 65.5) - P(Y < 45.5)$

$= P(Y < 65.5) - [1 - P(Y > 45.5)]$

$= P(Y < 65.5) + P(Y > 45.5) - 1$

$= P\left(z < \frac{65.5-60}{\sqrt{51}}\right) + P\left(> \frac{45.5-60}{\sqrt{51}}\right) - 1$

$= \Phi(0.770154...) + \Phi(-2.030406...) - 1$

$= 0.77934... + 0.97884 - 1$

$= 0.7582$

c) $X \sim B(20, 0.15)$

$H_0 : p = 0.15$

$H_1 : p \neq 0.15$, WHERE p REPRESENTS THE PROPORTION OF MISSED APPOINTMENTS IN GENERAL

CRITICAL REGION AT 10% SIGNIFICANCE IS REQUIRED, AT 5% EACH TAIL

- $P(X \leqslant 0) = 0.0388 = 3.88\% < 5\%$
- $P(X \leqslant 1) = 0.1756 = 17.56\% > 5\%$

⋮

- $P(X \geqslant 6) = 1 - P(X \leqslant 5) = 1 - 0.9327 = 0.0673 = 6.73\% > 5\%$
- $P(X \geqslant 7) = 1 - P(X \leqslant 6) = 1 - 0.9781 = 0.0219 = 2.19\% < 5\%$

$\therefore$ CRITICAL REGION $= \{0\} \cup \{7, 8, 9, \ldots 20\}$

d) LOOKING AT ABOVE "TABLES"

- $P(X \leqslant 0) = 3.88\%$ ← CLOSER
- $P(X \leqslant 1) = 17.56\%$

⋮

- $P(X \geqslant 6) = 6.73\%$ ← CLOSER
- $P(X \geqslant 7) = 2.19\%$

$\therefore C.R = \{0\} \cup \{6, 7, 8, \ldots, 20\}$

Question 2 (**+)

A teacher is investigating the students' method of getting back home and is told that 15% of the students gets back home by car.

He decides to investigate this fact further and decides to use a random sample of 36 students across all the school's year groups.

Stating your hypotheses clearly, find the critical region …

a) … for his test at the 6% level of significance.

b) … for a similar test where the probability of rejecting at either tail must be as close as possible to 3%.

$\boxed{x \le 1 \text{ or } x \ge 11}$, $\boxed{x \le 1 \text{ or } x \ge 10}$

a) X = NUMBER OF STUDENTS WHICH GET DRIVEN BACK HOME
$X \sim B(36, 0.15)$

SETTING HYPOTHESES

$H_0 : p = 0.15$

$H_1 : p \neq 0.15$, WHERE p IS THE PROPORTION OF STUDENTS WHO GET DRIVEN IN GENERAL

CRITICAL REGION REQUIRED, AT 6%, TWO TAILED

FROM CALCULATOR

CRITICAL

$P(X \le 1) = 0.0212 = 2.12\% < 3\%$

$P(X \le 2) = 0.0776 = 7.76\% > 3\%$

$\vdots$

$P(X \ge 10) = 1 - P(X \le 9) = 1 - 0.9649 = 0.0351 = 3.51\% > 3\%$

CRITICAL

$P(X \ge 11) = 1 - P(X \le 10) = 1 - 0.9863 = 0.0137 = 1.37\% < 3\%$

$\therefore \text{C.R} = \{0, 1\} \cup \{11, 12, 13, \dots, 36\}$

b) IF WE REJECT AS CLOSE AS POSSIBLE TO 3% IN EACH TAIL

2.12% IS CLOSER TO 3% THAN 7.76%

BUT

3.51% IS CLOSER TO 3% THAN 1.37%

$\therefore \text{C.R} = \{0, 1\} \cup \{10, 11, 12, \dots, 36\}$

Question 3 (**+)

A teacher is investigating the students' after school activities and is told that 1 in 20 students attends martial arts classes.

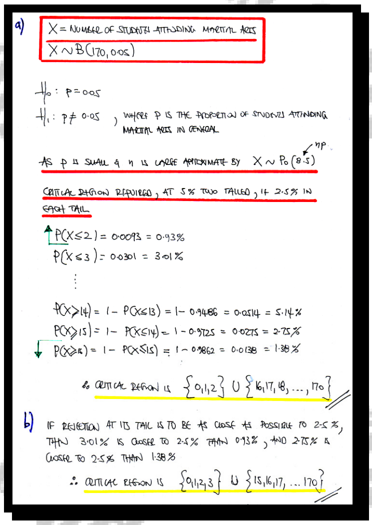
He decides to investigate this fact further and decides to use a random sample of 170 students across all the school's year groups.

By using a distributional approximation, and stating your hypotheses clearly, find the critical region …

a) … for his test at the 5% level of significance.

b) … for a similar test where the probability of rejecting at either tail must be as close as possible to 2.5% .

$\boxed{\text{FS1C}}$, $\boxed{x \le 2 \ \text{ or } \ x \ge 16}$, $\boxed{x \le 3 \ \text{ or } \ x \ge 15}$



a) X = NUMBER OF STUDENTS ATTENDING MARTIAL ARTS

$X \sim B(170, 0.05)$

$H_0: p = 0.05$

$H_1: p \neq 0.05$, WHERE p IS THE PROPORTION OF STUDENTS ATTENDING MARTIAL ARTS IN GENERAL

AS p IS SMALL & n IS LARGE APPROXIMATE BY $X \sim Po(\overset{np}{8.5})$

CRITICAL REGION REQUIRED, AT 5% TWO TAILED, IE 2.5% IN EACH TAIL

$P(X \le 2) = 0.0093 = 0.93\%$

$P(X \le 3) = 0.0301 = 3.01\%$

$\vdots$

$P(X \ge 14) = 1 - P(X \le 13) = 1 - 0.9486 = 0.0514 = 5.14\%$

$P(X \ge 15) = 1 - P(X \le 14) = 1 - 0.9725 = 0.0275 = 2.75\%$

$P(X \ge 16) = 1 - P(X \le 15) = 1 - 0.9862 = 0.0138 = 1.38\%$

$\therefore$ CRITICAL REGION IS $\{0, 1, 2\} \cup \{16, 17, 18, \dots, 170\}$

b) IF REJECTION AT ITS TAIL IS TO BE AS CLOSE AS POSSIBLE TO 2.5%, THEN 3.01% IS CLOSER TO 2.5% THAN 0.93%, AND 2.75% IS CLOSER TO 2.5% THAN 1.38%

$\therefore$ CRITICAL REGION IS $\{0, 1, 2, 3\} \cup \{15, 16, 17, \dots 170\}$

Question 4 (**+)

Hooks produced in a factory are packed in boxes of sixty. It is thought that 5% of the hooks produced are defective.

The main production machine is replaced and the makers of the new machine claim that their machine will produce less defective hooks.

A box from the production of the new machine is inspected and is found to contain one defective hook.

a) Test at the 10% level of significance the claim made by the makers of the new machine.

b) If the test was carried out at the 5% level of significance find the critical region for the test.

, $\boxed{\text{not significant evidence, } 19.16\% > 10\%}$, $\boxed{\text{C.R.} = \{0\}}$

a) $\boxed{\begin{array}{l} X = \text{NUMBER OF DEFECTIVE HOOKS} \\ X \sim B(60, 0.05) \end{array}}$

$H_0 : p = 0.05$
$H_1 : p < 0.05$, WHERE p REPRESENTS THE PROPORTION OF ALL THE DEFECTIVE HOOKS FROM THE PRODUCTION

TESTING AT 10% LEVEL OF SIGNIFICANCE ON THE BASIS THAT $x=1$

$P(X \leqslant 1) = $ CALCULATOR OR $P(X=0) + P(X=1)$
$= 0.1916$
$= 19.16\%$
$> 10\%$

THERE IS NO SIGNIFICANT EVIDENCE TO SUPPORT THE CLAIM MADE BY THE MAKERS OF THE NEW MACHINERY
NO SUFFICIENT EVIDENCE TO REJECT H_0

b) LOOKING AT THE "BOTTOM TAIL" AT 5%

$P(X \leqslant 0) = 0.0461 = 4.61\% < 5\%$
$P(X \leqslant 1) = 0.1916 = 19.16\% > 5\%$

$\therefore$ CRITICAL REGION $= \{0\}$

Question 5 (**+)

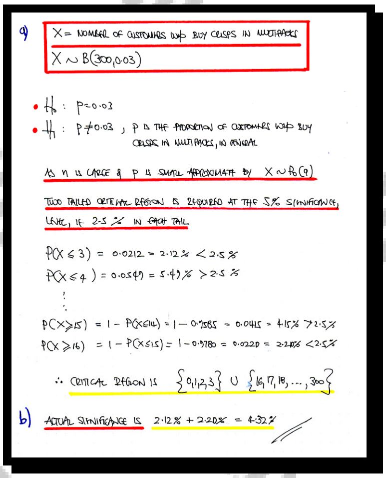
The manager of a supermarket believes that of the customers who buy crisps, 3% buy them in "Mega Packs".

He decides to test his belief by recording how many of the next 300 customers who buy crisps, bought them in "Mega Packs".

a) Using a distributional approximation, find the critical region to test the manager's belief, at the 5% level of significance.

b) State the actual significance level for a test using the critical region of part (**a**).

$\boxed{\text{C.R.} = \{0,1,2,3\} \cup \{16,17,18,\ldots,300\}}$, $\boxed{4.32\%}$



a) X = NUMBER OF CUSTOMERS WHO BUY CRISPS IN MULTIPACKS

$X \sim B(300, 0.03)$

- H_0 : $p = 0.03$
- H_1 : $p \neq 0.03$, p IS THE PROPORTION OF CUSTOMERS WHO BUY CRISPS IN MULTIPACKS, IN GENERAL

AS n IS LARGE & p IS SMALL APPROXIMATE BY $X \sim Po(9)$

TWO TAILED CRITICAL REGION IS REQUIRED AT THE 5% SIGNIFICANCE LEVEL, IE 2.5 % IN EACH TAIL

$P(X \leq 3) = 0.0212 = 2.12\% < 2.5\%$

$P(X \leq 4) = 0.0549 = 5.49\% > 2.5\%$

$\vdots$

$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9585 = 0.0415 = 4.15\% > 2.5\%$

$P(X \geq 16) = 1 - P(X \leq 15) = 1 - 0.9780 = 0.0220 = 2.20\% < 2.5\%$

$\therefore$ CRITICAL REGION IS $\{0,1,2,3\} \cup \{16,17,18,\ldots,300\}$

b) ACTUAL SIGNIFICANCE IS $2.12\% + 2.20\% = 4.32\%$

POISSON DISTRIBUTION HYPOTHESIS TESTING

Question 1

Bacteria are randomly distributed in the water of a lake at the rate of 7 per litre.

A new factory opens next to the lake and the local residents feel that the factory is polluting the lake. To test this claim, a random sample of half a litre of water from the lake is examined and found to contain 8 bacteria.

Test at the 5% level of significance, whether there is evidence to support the claim of the local residents.

$\boxed{}$, $\boxed{\text{significant evidence, } 2.67\% < 5\%}$

THE RATE IS 7 BACTERIA PER LITRE — ADJUST TO HALF LITRE, SO 3.5 BACTERIA PER HALF LITRE

X = NO OF BACTERIA PER HALF LITRE

$X \sim Po(3.5)$

$H_0 : \lambda = 3.5$

$H_1 : \lambda > 3.5$, WHERE λ IS THE RATE OF BACTERIA IN GENERAL (NOT IN THE HALF LITRE WE COLLECTED)

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 8$

$$P(X \geqslant 8) = 1 - P(X \leqslant 7)$$
$$= 1 - 0.97326\ldots$$
$$= 0.026738\ldots$$
$$= 2.67\%$$
$$< 5\%$$

THERE IS SIGNIFICANT EVIDENCE TO SUPPORT THE CLAIM MADE BY THE LOCAL RESIDENTS — SUFFICIENT EVIDENCE TO REJECT H_0

POISSON DISTRIBUTION HYPOTHESIS TESTING

EXAM QUESTIONS

Question 1 (**)

A company advertises for a summer job every year.

It has been established over a long period of years that the number of applicants per year follows a Poisson distribution with mean 12.

This year there were 19 applicants for this summer job.

Test, at the 1% level of significance, whether there is evidence of an increase of the mean number of the applicants for the job.

$\boxed{}$, $\boxed{\text{not significant, } 3.74\% > 1\%}$

$\boxed{\begin{array}{l} X = \text{NO OF APPLICANTS PER YEAR} \\ X \sim Po(12) \end{array}}$

SETTING UP SUITABLE HYPOTHESES

$H_0: \lambda = 12$

$H_1: \lambda > 12$

WHERE λ IS THE RATE OF APPLICANTS PER YEAR, IN GENERAL

TESTING AT THE 1% LEVEL OF SIGNIFICANCE ON THE BASIS THAT $x = 19$

$$\begin{aligned} P(X \geqslant 19) &= 1 - P(X \leqslant 18) \\ &\quad \ldots\text{tables/calculator}\ldots \\ &= 1 - 0.9626 \\ &= 0.0376 \\ &= 3.76\% > 1\% \end{aligned}$$

THERE IS NO SIGNIFICANT EVIDENCE, AT 1% LEVEL, THAT THERE HAS BEEN AN INCREASE IN THE NUMBER OF APPLICANTS PER YEAR.

NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 2 (**+)

Minor imperfections occur in the cloth manufacture in a certain factory at the rate of 1.2 per square metre of cloth.

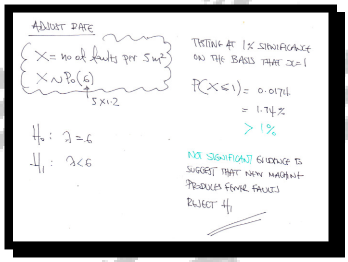
The factory owner buys a new cloth manufacturing machine.

The makers of the new machine claim that their product will reduce these minor imperfections.

A random 5 square metre piece of cloth produced by the new machine is examined and found to contain 1 minor imperfection.

Test at the 1% level of significance, whether there is evidence to support the claim of the makers of the new machine.

$\boxed{\text{not significant evidence, } 1.74\% > 1\%}$



Question 3 (***)

George works in a shoe shop on Saturdays.

On average he serves 8 customers every hour.

a) Find the probability that, in a given hour on a Saturday, George will serve more than 10 customers.

b) Find the probability that, in a given three hour interval on a Saturday, George will serve exactly 24 customers.

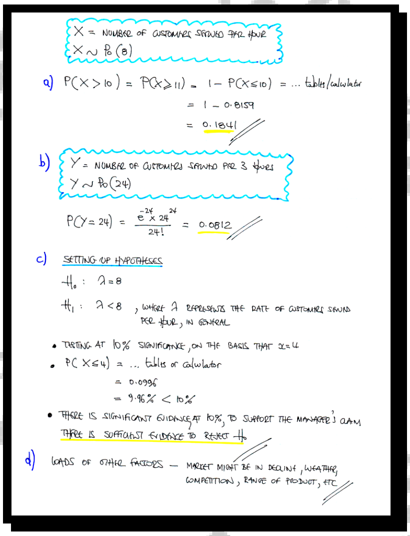
The manager of the shop claims that George has recently become complaisant and serves fewer customers.

During a given hour on a particular Saturday George serves 4 customers.

c) Test at the 10% level of significance, whether there is evidence to support the manager's claim.

d) Comment on whether other factors might affect the reliability of the test.

$\boxed{}$, $\boxed{0.1841}$, $\boxed{0.0812}$, $\boxed{\text{significant evidence, } 9.96\% < 10\%}$, $\boxed{\text{not reliable; maybe business is not brisk}}$


X = NUMBER OF CUSTOMERS SERVED PER HOUR
$X \sim Po(8)$

a) $P(X>10) = P(X \geqslant 11) = 1 - P(X \leqslant 10) = \ldots$ tables/calculator
$= 1 - 0.8159$
$= 0.1841$

b) Y = NUMBER OF CUSTOMERS SERVED PER 3 HOURS
$Y \sim Po(24)$

$P(Y=24) = \dfrac{e^{-24} \times 24^{24}}{24!} = 0.0812$

c) SETTING UP HYPOTHESES

$H_0 : \lambda = 8$

$H_1 : \lambda < 8$, WHERE λ REPRESENTS THE RATE OF CUSTOMERS SERVED PER HOUR, IN GENERAL

- TESTING AT 10% SIGNIFICANCE, ON THE BASIS THAT $x=4$
- $P(X \leqslant 4) = \ldots$ tables or calculator
$= 0.0996$
$= 9.96\% < 10\%$
- THERE IS SIGNIFICANT EVIDENCE, AT 10%, TO SUPPORT THE MANAGER'S CLAIM
THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

d) LOADS OF OTHER FACTORS – MARKET MIGHT BE IN DECLINE, WEATHER, COMPETITION, RANGE OF PRODUCT, ETC

Question 4 (***+)

Minor imperfections occur in the cloth manufacture in a certain factory at the rate of 1.6 per square metre of cloth.

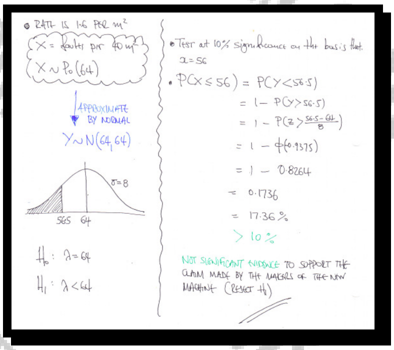
The factory owner buys a new cloth manufacturing machine.

The makers of the new machine claim that their product will reduce the number of these minor imperfections.

A random 40 square metre piece of cloth produced by the new machine is examined and found to contain 56 minor imperfections.

Test at the 10% level of significance, whether there is evidence to support the claim of the makers of the new machine.

$\boxed{\text{not significant evidence, } 17.36\% > 10\%}$


• RATE IS 1.6 PER m^2

X = faults per $40\ m^2$

$X \sim Po(64)$

APPROXIMATE BY NORMAL

$Y \sim N(64, 64)$

$\sigma = 8$

56.5 64

$H_0: \lambda = 64$

$H_1: \lambda < 64$

• Test at 10% significance on the basis that $x = 56$

• $P(X \leq 56) = P(Y < 56.5)$

$= 1 - P(Y > 56.5)$

$= 1 - P\left(Z > \frac{56.5 - 64}{8}\right)$

$= 1 - \Phi(0.9375)$

$= 1 - 0.8264$

$= 0.1736$

$= 17.36\%$

$> 10\%$

NOT SIGNIFICANT EVIDENCE TO SUPPORT THE CLAIM MADE BY THE MAKERS OF THE NEW MACHINE (REJECT H_1)

Question 5 (***+)

It is thought that calls arrive at a company telephone switchboard at the steady rate of 7 per minute.

a) Find the probability that, in a given minute, there will be more than 5 but at most 10 calls arriving the company's telephone switchboard.

A 5 minute interval is divided into 10 equal 30 second intervals.

b) Find the probability that there will be at least one 30 second interval without a single call arriving the company's telephone switchboard.

The telephone operator claims that the rate of calls has risen recently, and asks her manager for a helper.

The manager investigates this matter and finds that in a randomly chosen minute, 13 calls reached the telephone switchboard.

c) Test at the 5% level of significance, whether there is evidence to support the telephone operator's claim.

$\boxed{}$, $\boxed{0.6008}$, $\boxed{0.2641}$, $\boxed{\text{significant evidence, } 2.7\% < 5\%}$

$\boxed{\begin{array}{l} X = \text{NO OF CALLS PER MINUTE} \\ X \sim Po(7) \end{array}}$

a) $P(\text{more than 5 but at most 10}) = P(5 < X \leqslant 10)$
$= P(6 \leqslant X \leqslant 10)$
$= P(X \leqslant 10) - P(X \leqslant 5)$
$= 0.9015 - 0.3007$
$= 0.6008$

b) READJUST THE RATE TO $\frac{1}{2}$ MINUTE

$\boxed{\begin{array}{l} Y = \text{NO OF CALLS PER } \tfrac{1}{2} \text{ MINUTE} \\ Y \sim Po(3.5) \end{array}}$

$P(Y=0) = \dfrac{e^{-3.5} \times 3.5^0}{0!} = 0.0302$

$\boxed{\begin{array}{l} W = \text{NO OF } \tfrac{1}{2} \text{ INTERVALS WITHOUT A CALL} \\ W \sim B(10, 0.0302) \end{array}}$

$P(W \geqslant 1) = 1 - P(W \leqslant 0) = 1 - P(W=0) = 1 - \binom{10}{0}(0.0302)^0(0.9698)^{10}$
$= 1 - 0.7359\ldots \simeq 0.2641$

c) SETTING SUITABLE HYPOTHESES

$H_0: \lambda = 7$
$H_1: \lambda > 7$

WHERE λ REPRESENTS THE AVERAGE RATE OF CALLS PER MINUTE, IN GENERAL

TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $x = 13$

$P(X \geqslant 13) = 1 - P(X \leqslant 12)$
$= 1 - 0.9730$
$= 0.0270$
$= 2.7\% < 5\%$

THERE IS SIGNIFICANT EVIDENCE, AT THE 5% LEVEL, TO SUPPORT THE TELEPHONE OPERATOR'S CLAIM

SUFFICIENT EVIDENCE TO REJECT H_0

Question 6 (****)

Since his retirement, Fred goes fishing Monday to Friday, for 3 hours on each of these 5 days. The number of fish he catches every hour follows a Poisson distribution with mean 2.5 .

a) Find the probability that Fred catches more than 9 fish on exactly 2 of the days, in a given 5 day fishing week.

Fred buys a new type of bait and decides to test whether there is any difference to the rate at which he catches fish. He tries his new bait by going fishing on a Sunday and ends up catching 16 fish in 4 hours.

b) Carry out a significance test, at the 5% level, stating your hypotheses clearly.

$\boxed{}$, $\boxed{0.234}$, $\boxed{\text{not significant evidence, } 4.87\% > 2.5\%}$

a) START BY DEFINING VARIABLES AND DISTRIBUTIONS

"FISH CATCHING" RATE $\Rightarrow$ 2.5 fish per hour

ADJUSTING THE RATE TO 3 HOURS

$\boxed{\begin{array}{l} X = \text{NO OF FISH CAUGHT PER 3 HOURS} \\ X \sim Po(7.5) \end{array}}$

- $P(X>9) = P(X \geq 10) = 1 - P(X \leq 9) = \ldots$ tables $\ldots$
  $= 1 - 0.7764 = 0.2236$

$\boxed{\begin{array}{l} Y = \text{NO OF DAYS (OUT OF 5), WHERE MORE THAN 9 FISH IS CAUGHT} \\ Y \sim B(5, 0.2236) \end{array}}$

- $P(Y=2) = \binom{5}{2}(0.2236)^2(0.7764)^3 = 0.234$ //

b) ADJUSTING THE RATE TO 4 HOURS — $4 \times 2.5 = 12.5$

$\boxed{\begin{array}{l} W = \text{NO OF FISH CAUGHT PER 4 HOURS} \\ W \sim Po(10) \end{array}}$

$H_0 : \lambda = 2.5 \ (\mu = 10)$

$H_1 : \lambda \neq 2.5 \ (\mu = 10)$

- TESTING AT 5% SIGNIFICANCE ON THE BASIS THAT $W = 16$
- $P(W \geq 16) = 1 - P(W \leq 15)$
  $= 1 - 0.9513$
  $= 0.0487 > 2.5\%$

THERE IS NO SIGNIFICANT EVIDENCE TO SUGGEST THAT THE "FISH CATCHING" RATE IS DIFFERENT — NOT SUFFICIENT EVIDENCE TO REJECT H_0 //

Question 7 (****)

The number of accidents occurring in a certain stretch of motorway is modelled by a Poisson distribution with mean of 0.7 per day. The traffic police introduce speed camera in an effort to reduce this rate. They decide to monitor this stretch of motorway for 60 days after the introduction of the cameras.

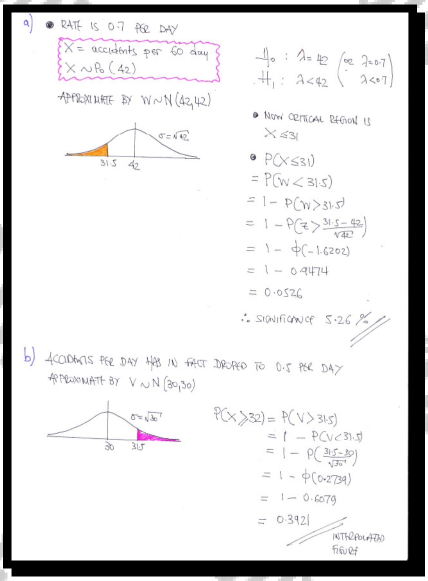
A hypothesis test is to be carried out afterwards. It is decided that the cameras would have had the desired effect if 31 or less accidents occur during these 60 days.

a) Determine the significance level for this test. State the hypotheses clearly.

It is now given that the introduction of cameras has in fact reduced the mean accident rate to 0.5 per day.

b) Find the probability that the traffic police will conclude that the cameras have not reduced the accident rate.

$\boxed{\text{significant level} = 5.26\%}$, $\boxed{0.3921}$



a) • RATE IS 0.7 PER DAY

X = accidents per 60 day

$X \sim Po(42)$

$H_0 : \lambda = 42 \quad (\text{or } \lambda = 0.7)$

$H_1 : \lambda < 42 \quad (\lambda < 0.7)$

APPROXIMATE BY $W \sim N(42, 42)$

$\sigma = \sqrt{42}$

31.5 42

• NOW CRITICAL REGION IS $X \leqslant 31$

• $P(X \leqslant 31)$

$= P(W < 31.5)$

$= 1 - P(W > 31.5)$

$= 1 - P\left(z > \frac{31.5 - 42}{\sqrt{42}}\right)$

$= 1 - \Phi(-1.6202)$

$= 1 - 0.9474$

$= 0.0526$

$\therefore$ SIGNIFICANCE 5.26 %

b) ACCIDENTS PER DAY HAS IN FACT DROPED TO 0.5 PER DAY

APPROXIMATE BY $V \sim N(30, 30)$

$\sigma = \sqrt{30}$

30 31.5

$P(X \geqslant 32) = P(V > 31.5)$

$= 1 - P(V < 31.5)$

$= 1 - P\left(\frac{31.5 - 30}{\sqrt{30}}\right)$

$= 1 - \Phi(0.2739)$

$= 1 - 0.6079$

$= 0.3921$

INTERPOLATED FIGURE

Question 8 (****+)

It is known that during the first hour of trading, customers arrive at a garden centre at the rate of 3 customers every 10 minutes.

a) State two conditions, so that a Poisson distribution could be used to model the number of arrivals in the garden centre.

b) State one reason as to why the Poisson model might not be suitable.

Using a suitable Poisson model, and ignoring the answer to part **(b)**, determine the probability that...

c) … exactly 4 customers arrive in the first 10 minutes of trading.

d) … exactly 8 customers arrive in the first 20 minutes of trading.

e) … exactly 4 customers arrive in the first 10 minutes of trading and a further 4 customers arrive in the next 10 minutes.

f) … the **first** customer will arrive 7 minutes after opening.

The first hour after opening is subdivided into 6 equal 10 minute intervals.

Calculate the probability that...

g) … exactly 4 customers arrive in **none** of these 10 minute intervals.

h) … there will be a **single** 10 minute interval, where **no** customers arrive to the garden centre.

The garden centre claims that 98% of its seeds will germinate. It was found that in a random sample of 125 seeds, 117 seeds germinated.

i) Use a suitable approximation, to test at the 1% level of significance, whether the garden centre overstates thr germination proportion of its seeds. You must state your hypotheses clearly in this part.

$\boxed{}$, $\boxed{0.1680}$, $\boxed{0.1033}$, $\boxed{0.0282}$, $\boxed{0.1225}$, $\boxed{0.3316}$, $\boxed{0.2314}$,

$\boxed{\text{significant evidence, } 0.42\% < 1\%}$

[solution overleaf]

a) • CUSTOMERS ARRIVE INDEPENDENTLY OF ONE ANOTHER
• CUSTOMERS ARRIVE AT A UNIFORM CONSTANT RATE PER UNIT TIME

b) CUSTOMERS ARE UNLIKELY TO ARRIVE "SINGLY", AS GARDEN CENTRES TEND TO BE VISITED BY COUPLES OR FAMILIES

c) X = NO OF CUSTOMER ARRIVALS PER 10 MINUTES
$X \sim Po(3)$

$$P(X=4) = \frac{e^{-3} \times 3^4}{4!} = 0.1680$$

d) Y = NO OF CUSTOMER ARRIVALS PER 20 MINUTES
$Y \sim Po(6)$

$$P(Y=8) = \frac{e^{-6} \times 6^8}{8!} = 0.1033$$

e) USING THE RESULT OF PART (c)

$$P(X=4) \times P(X=4) = \left(\frac{e^{-3} \times 3^4}{4!}\right)^2 = 0.0282$$

f) EITHER

V = NO OF CUSTOMERS PER MINUTE
$V \sim Po(0.3)$

$$P(V=0) = \frac{e^{-0.3} \times 0.3^0}{0!} = e^{-0.3}$$

$\therefore$ REQUIRED PROBABILITY

$$(e^{-0.3})^7 = e^{-2.1} \approx 0.122$$

OR

W = NO OF CUSTOMERS PER 7 MINUTES
$W \sim Po(2.1)$

$$P(W=0) = \frac{e^{-2.1} \times 2.1^0}{0!} \approx 0.1225$$

g) MODEL AS FOLLOWS, USING VARIABLE X, AS PREVIOUSLY DEFINED

$$P(X=4) = \frac{e^{-3} \times 3^4}{4!}$$

$$\text{REQUIRED PROBABILITY} = \left(1 - \frac{e^{-3} \times 3^4}{4!}\right)^6 = 0.3316$$

h) USING X AGAIN

$$P(X=0) = \frac{e^{-3} \times 0^3}{0!} = e^{-3}$$

REQUIRED PROBABILITY IS $(1-e^{-3})^5 \times e^{-3} \times 6$ WAYS

OR DEFINE A BINOMIAL

U = NO OF INTERVALS WITHOUT CUSTOMERS
$U \sim B(6, e^{-3})$

$$P(U=1) = \binom{6}{1} \times (e^{-3})^1 \times (1-e^{-3})^5 \approx 0.2314$$

i) G = NO OF GERMINATING SEEDS
G' = NO OF NON GERMINATING SEEDS

$G \sim B(125, 0.98)$ AND $G' \sim B(\overset{n}{125}, \overset{p}{0.02})$

$\downarrow$ APPROX BY POISSON

$G' \sim Po(\overset{np}{2.5})$

• TO TEST THE HYPOTHESIS WITH G $\quad P(G \leqslant 117)$ (117 GERMINATED)
• TO TEST THE HYPOTHESIS WITH G' $\quad P(G' \geqslant 8)$ (8 DIDN'T GERMINATE)

$\left.\begin{array}{l} H_0: p = 0.02 \\ H_1: p > 0.02 \end{array}\right\}$ WHERE p IS THE PROPORTION OF NON GERMINATING SEEDS IN THE ENTIRE POPULATION

• $P(G' \geqslant 8) = 1 - P(G' \leqslant 7)$
$= 1 - 0.9958$
$= 0.0042$
$= 0.42\%$
$< 1\%$

THERE IS SIGNIFICANT EVIDENCE THAT THE GARDEN CENTRE OVERSTATES THE GERMINATION PROPORTION OF ITS SEEDS – SUFFICIENT EVIDENCE TO REJECT H_0

NOTE IT IS FINE TO PUT HYPOTHESES AS

$\left.\begin{array}{l} p = 0.98 \\ p < 0.98 \end{array}\right\}$ WHERE p IS THE PROPORTION OF GERMINATING SEEDS

& THE REST IDENTICAL AS ABOVE

POISSON DISTRIBUTION CRITICAL REGIONS

Question 1

The new chief of the traffic police was told that in a busy stretch of motorway, under his jurisdiction, accidents occur at the rate of 7 per month.

The new chief wants to know whether this figure is up to date and decides to test it to see if this rate is different from 7 accidents per month.

a) Find the critical region for this test.

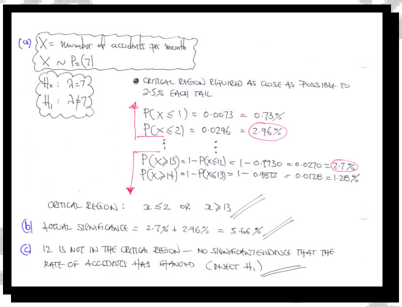
The probability of rejecting at either tail must be as close as possible to 2.5%.

b) State the actual significance level for a test using the critical region of part **(a)**.

In the month following the new chief's appointment there were 12 accidents in that stretch of motorway.

c) Complete the test.

$\boxed{\text{C.R.} = \{0,1,2\} \cup \{13,14,15,\ldots\}}$, $\boxed{5.66\%}$, $\boxed{\text{not significant evidence, } 12 \text{ is not in C.R.}}$



(a) $X =$ number of accidents per month

$X \sim Po(7)$

$H_0 : \lambda = 7$

$H_1 : \lambda \neq 7$

• CRITICAL REGION REQUIRED AS CLOSE AS POSSIBLE TO 2.5% EACH TAIL

$P(X \leq 1) = 0.0073 = 0.73\%$

$P(X \leq 2) = 0.0296 = 2.96\%$

$\vdots$

$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.9730 = 0.0270 = 2.7\%$

$P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.9872 = 0.0128 = 1.28\%$

CRITICAL REGION: $x \leq 2$ OR $x \geq 13$

(b) ACTUAL SIGNIFICANCE $= 2.7\% + 2.96\% = 5.66\%$

(c) 12 IS NOT IN THE CRITICAL REGION — NO SIGNIFICANT EVIDENCE THAT THE RATE OF ACCIDENTS HAS CHANGED (REJECT H_1)

Question 2

A test statistic has distribution $\text{Po}(\lambda)$.

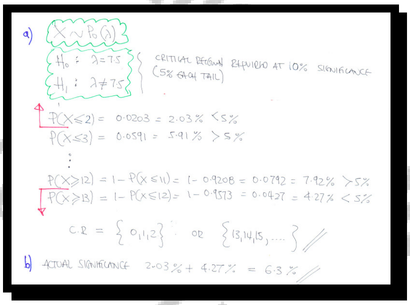
a) Given that

$$H_0 : \lambda = 7.5 , \quad H_1 : \lambda \neq 7.5 ,$$

find the critical region for the test statistic at the 10% significance level.

b) State the actual level of significance in a test using this critical region.

$\boxed{\{0,1,2\} \cup \{13,14,15,16,...\}}$, $\boxed{6.3\%}$


a) $X \sim Po(\lambda)$

$H_0 : \lambda = 7.5$
$H_1 : \lambda \neq 7.5$ CRITICAL REGION REQUIRED AT 10% SIGNIFICANCE (5% EACH TAIL)

$P(X \leq 2) = 0.0203 = 2.03\% < 5\%$
$P(X \leq 3) = 0.0591 = 5.91\% > 5\%$
$\vdots$
$P(X \geq 12) = 1 - P(X \leq 11) = 1 - 0.9208 = 0.0792 = 7.92\% > 5\%$
$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.9573 = 0.0427 = 4.27\% < 5\%$

$C.R = \{0,1,2\}$ OR $\{13,14,15,....\}$ //

b) ACTUAL SIGNIFICANCE $2.03\% + 4.27\% = 6.3\%$ //

Question 3

A test statistic has distribution $\text{Po}(\lambda)$.

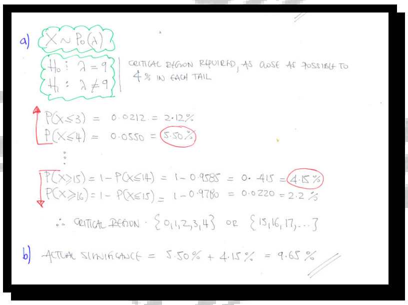
a) Given that

$$\text{H}_0 : \lambda = 9, \quad \text{H}_1 : \lambda \neq 9,$$

find the critical region for the test statistic such that the probability of rejecting in each tail is as close as possible to 4%.

b) State the probability of incorrectly rejecting H_0 using this critical region.

$\boxed{\{0,1,2,3,4\} \cup \{15,16,17,18,\ldots\}}$, $\boxed{9.65\% \text{ or } 0.0965}$



a) $X \sim \text{Po}(\lambda)$

$\text{H}_0 : \lambda = 9$
$\text{H}_1 : \lambda \neq 9$

CRITICAL REGION REQUIRED, AS CLOSE AS POSSIBLE TO 4% IN EACH TAIL

$P(X \leq 3) = 0.0212 = 2.12\%$
$P(X \leq 4) = 0.0550 = 5.50\%$

$\vdots$

$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9585 = 0.415 = 4.15\%$
$P(X \geq 16) = 1 - P(X \leq 15) = 1 - 0.9780 = 0.0220 = 2.2\%$

$\therefore$ CRITICAL REGION $\{0,1,2,3,4\}$ OR $\{15,16,17,\ldots\}$

b) ACTUAL SIGNIFICANCE $= 5.50\% + 4.15\% = 9.65\%$

POISSON DISTRIBUTION CRITICAL REGIONS

EXAM QUESTIONS

Question 1 (**+)

During Saturday afternoons, customers are known to walk into a certain clothes shop at the rate of 8 every 10 minutes.

The new shop manager wants to find if that rate has changed since he took over.

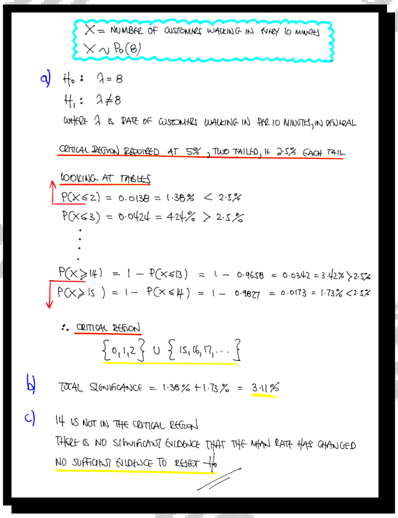
a) Find the critical region, at the 5% level of significance, to investigate whether the rate of 8 customers walking into the shop every 10 minutes has changed.

b) State the actual significance level for a test using the critical region of part (**a**).

During a Saturday afternoon, 14 customers walked into the shop in a random 10 minute interval.

c) Complete the test.

$\boxed{}$, $\boxed{\text{C.R.} = \{0,1,2\} \cup \{15,16,17,\ldots\}}$, $\boxed{3.11\%}$, $\boxed{\text{not significant evidence, 14 is not in C.R.}}$



X = NUMBER OF CUSTOMERS WALKING IN EVERY 10 MINUTES
$X \sim Po(8)$

a) $H_0: \lambda = 8$
$H_1: \lambda \neq 8$
WHERE λ IS RATE OF CUSTOMERS WALKING IN PER 10 MINUTES, IN GENERAL

CRITICAL REGION REQUIRED AT 5%, TWO TAILED, IE 2.5% EACH TAIL

LOOKING AT TABLES

$P(X \leqslant 2) = 0.0138 = 1.38\% < 2.5\%$
$P(X \leqslant 3) = 0.0424 = 4.24\% > 2.5\%$
$\vdots$
$P(X \geqslant 14) = 1 - P(X \leqslant 13) = 1 - 0.9658 = 0.0342 = 3.42\% > 2.5\%$
$P(X \geqslant 15) = 1 - P(X \leqslant 14) = 1 - 0.9827 = 0.0173 = 1.73\% < 2.5\%$

$\therefore$ CRITICAL REGION
$\{0,1,2\} \cup \{15,16,17,\ldots\}$

b) TOTAL SIGNIFICANCE $= 1.38\% + 1.73\% = 3.11\%$

c) 14 IS NOT IN THE CRITICAL REGION
THERE IS NO SIGNIFICANT EVIDENCE THAT THE MEAN RATE HAS CHANGED
NO SUFFICIENT EVIDENCE TO REJECT H_0

Question 2 (***+)

From a Poisson distribution with parameter λ, a single observation w is taken and is to be used to test, at the 5% level of significance,

$$\mathrm{H}_0 : \lambda = k \qquad \text{against} \qquad \mathrm{H}_1 : \lambda \neq k,$$

where k is a positive integer.

Find the actual significance for the test if the critical region is

$$W \le 1 \cup W \ge 12.$$

$\boxed{}$, $\boxed{3.75\%}$

- AS THE TEST IS TWO TAILED, THE SIGNIFICANCE MUST BE 2.5% IN EACH TAIL
- LOOKING AT THE POISSON TABLES
  - "BOTTOM END"

    $W = 0, 1$ THIS OCCURS FOR INTEGER VALUES OF λ IF $k = 6, 7$
  - "TOP END"

    $W = 12, 13, 14, \dots$ THIS OCCURS FOR INTEGER VALUES OF λ IF $k = 6$ (ONLY)
- HENCE $k = 6$ AND BY USING TABLES STILL

  $$P(W \le 1) = \dots \text{tables} = 0.0174 = 1.74\%$$
  $$P(W \ge 12) = \dots \text{tables} = 1 - P(W \le 11) = 1 - 0.9799 = 0.0201 = 2.01\%$$
- ACTUAL SIGNIFICANCE IS $1.74\% + 2.01\% = 3.75\%$

Question 3 (***+)

The proportion of customers who buy alcoholic drinks with their shopping from a certain supermarket is 3%.

The new manager of the supermarket believes that this proportion is different from 3%.

He decides to test his belief using a random sample of 300 customers.

a) Find the critical region at the 5% level of significance to test the manager's belief.

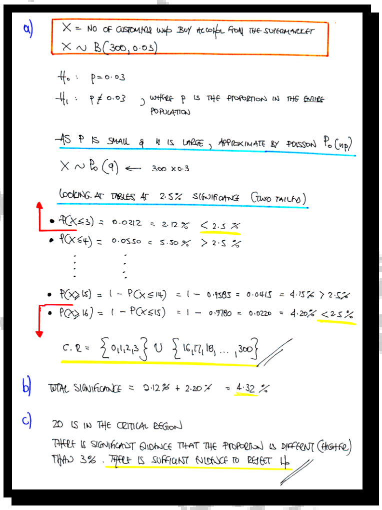
b) State the actual significance for a test using the critical region of part **(a)**.

Twenty shoppers in the sample bought alcoholic drinks.

c) Complete the test.

$\boxed{}$, $\boxed{\text{C.R.} = \{0,1,2,3\} \cup \{16,16,18,\ldots,300\}}$, $\boxed{4.32\%}$,

$\boxed{\text{significant evidence, 20 is in C.R.}}$



a) X = NO OF CUSTOMERS WHO BUY ALCOHOL FROM THE SUPERMARKET

$X \sim B(300, 0.03)$

$H_0: p = 0.03$

$H_1: p \neq 0.03$, WHERE p IS THE PROPORTION IN THE ENTIRE POPULATION

AS p IS SMALL & n IS LARGE, APPROXIMATE BY POISSON $Po(np)$

$X \sim Po(9) \leftarrow 300 \times 0.3$

LOOKING AT TABLES AT 2.5% SIGNIFICANCE (TWO TAILED)

- $P(X \leq 3) = 0.0212 = 2.12\% < 2.5\%$
- $P(X \leq 4) = 0.0550 = 5.50\% > 2.5\%$

$\vdots \qquad \vdots$

- $P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9585 = 0.0415 = 4.15\% > 2.5\%$
- $P(X \geq 16) = 1 - P(X \leq 15) = 1 - 0.9780 = 0.0220 = 4.20\% < 2.5\%$

$C.R. = \{0,1,2,3\} \cup \{16,17,18,\ldots,300\}$

b) TOTAL SIGNIFICANCE $= 2.12\% + 2.20\% = 4.32\%$

c) 20 IS IN THE CRITICAL REGION

THERE IS SIGNIFICANT EVIDENCE THAT THE PROPORTION IS DIFFERENT (HIGHER) THAN 3%. THERE IS SUFFICIENT EVIDENCE TO REJECT H_0