Experimental design CW2

Wednesday, 23 June 2021 15:37

Name: Rohan Killedar
Student Number: 100954251

Part 1

1. In the 2014 referendum on the independence of Scotland from the UK, approx. 44% of Scotland's population was pro-independence. Suppose we now interview a random sample of 100 people. What is the probability that 50 or more of them are pro-independence? (assuming that the Scotland population preference didn't change since 2014) *[20 marks]*
tip: use normal approximation of binomial

Given : $p = 0.44$ & $n = 100$

$P(x \geq 50)$

Lets calculate mean of given binomial distribution

$\mu = np = 100 \times 0.44 = 44$

$\sigma^2 = np(1-p) = 44(1-0.44)$
$= 44 \times 0.56$
$= 24.64$
$\sigma = \sqrt{24.64} = 4.96$

Now using normal approximation

$P(x \geq 50) \approx P\left(z \geq \frac{50-44}{4.96}\right)$

$= P(z \geq 1.20) = 1 - P(z < 1.19)$
$= 1 - 0.88298$
$= 0.117$
≈ 0.12

So, the probability that 50 or more are pro-independance is 12 %

2. Royal Holloway conduct a study on the lifestyle of their students. They want to know about the average weekly hours of physical exercise of their students. From a sample of 30 students from Computer Science, they obtain a sample mean of $\bar{x}_{CS} = 1.5$ hours of exercise per week, with a sample standard deviation of $s_{CS} = 0.1$. From a sample of 30 students from Biology, they obtain a mean of $\bar{x}_{Bio} = 1.6$, with a sample standard deviation of $s_{Bio} = 0.15$.
   a) Identify and apply a suitable two-sample, two-sided test to establish whether there is a statistically significant difference between the mean weekly exercise hours of Computer Science and Biology students. Consider a significance level of $\alpha = 0.05$. (tip: use $df = 29$). *[14 marks]*
   b) Compute a 95% confidence interval for $\bar{x}_{CS} - \bar{x}_{Bio}$. *[9 marks]*
   c) Is the difference practically significant? Justify your answer *[7 marks]*

Given : $n_{CS} = 30$ $n_{Bio} = 30$
$\bar{x}_{CS} = 1.5$ $\bar{x}_{Bio} = 1.6$

$S_{cs} = 0.1 \qquad S_{Bio} = 0.15$

a) $\alpha = 0.05 \qquad df = n-1 = 29$

$$\text{Test statistic} = (\bar{x}_{cs} - \bar{x}_{Bio}) \Big/ \sqrt{\frac{s_{Bio}^2}{n_{Bio}} + \frac{S_{cs}^2}{n_{cs}}}$$

$$= (1.5 - 1.6) \Big/ \sqrt{\frac{(0.15)^2}{30} + \frac{(0.1)^2}{30}}$$

$$= -0.1 \,/\, 0.0329$$
$$= -3.0395$$

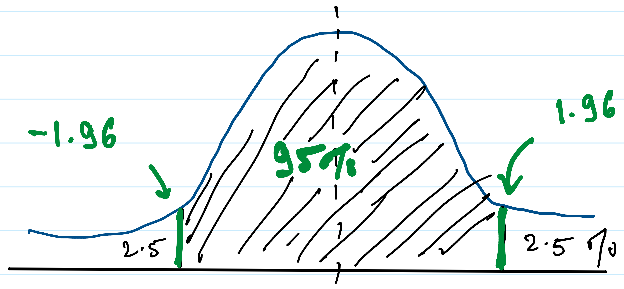
So, using t-table and test statistic we find p-value i.e. 0.01

Hence, statistically significant ($p < 0.05$)

b) for 95 % confidence interval

$$\bar{x}_{cs} - \bar{x}_{Bio} = -0.1$$

$$\sigma_{cs-Bio} = 0.0329$$



$$\mu_{\bar{x}_{cs} - \bar{x}_{Bio}}$$

$\therefore$ Confidence interval

$$= \left[\mu_{\bar{x}_{cs} - \bar{x}_{Bio}} \pm Z_{\alpha/2} \times \sigma_{cs-bio} \right]$$

$$= [-0.1 \pm 1.96 \times 0.0329]$$

$$= [-0.1 \pm 0.06448]$$

$= [-0.16448, -0.03552]$

c) Practically significance

The result are not pratically significant as the confidence interval and mean difference found are not really much when considered in terms of time spent doing exercise.
The mean difference of 0.1 is not much when you realise its 0.1 hour over a week time is just 6 minutes, that practically wont make much difference in terms of exercise.

3. Bob, an AI researcher, wants to prove that his image classification method works better than the state-of-the art. For this purpose, he takes 10 independent samples of images and for each sample, he performs a statistical test at level $\alpha = 0.05$ to establish if his method significantly outperforms the state-of-the-art. He finds that only 4 out of 10 tests successfully reject the null hypothesis (that Bob's method is no better than state-of-the-art). Bob publishes his work reporting only the test with the smallest p-value, claiming a significant improvement by his method.
   a) Is Bob's claim correct? If not, what technique should he use to rectify the results? Justify your answer *[12 marks]*
   b) what is the distribution, and its parameters, of the variable X that counts the number of tests for which Bob correctly fails to reject the null hypothesis (i.e., for which the null is not rejected when the null is true)? *[12 marks]*
   c) What's the expected value and variance of the above X? *[6 marks]*

a) This is the case of p-hacking where Bob only published result of smallest p-value but with significance level $\alpha = 5\%$, the chance of false positive in 10 sample is 0.5 sample, which perhaps can be the sample that Bob chose.

To fix this, he can use Bonferroni's correction

When performing 10 test, He need to adjust the overall significance level using

$1 - (1 - \alpha/m)^m$

where m is no. of tests

$\therefore$ overall significance $= 1 - (1 - 0.05/10)^{10}$
$= 0.0488$

Now, using the overall significance, Bob should determine whether his calculated p-values reject the null hypothesis or not.

b) As the random variable X takes 2 values i.e. success or failure to reject the null hypothesis and each test is independent of each other. Each test wrongly rejecting the null can be seen as a Bernoulli variable, with probability α. So, the total number of test that wrongly reject the null is the sum of iid Bernoulli variables hence a binomial with parameter $n = 10$ and $p = \alpha = 0.05$.

c) Here, the number of independent samples $n = 10$

We can take the value of Random Variable X as 1 for success and -1 for failure

$$\text{Expected Value} = \sum_{x = X} x \cdot p(x)$$

$$= 1 \times 0.95 + -1 \times 0.05$$

$$= 0.9$$

$$\text{variance } s^2 = np(1-p) = 10 \times 0.95 \times 0.5$$

$$= 4.75$$