RUNNING HEAD: SHORTENED TITLE: EPIDEMIOLOGY

Title of the Paper: Epidemiology for health and medical science

Author

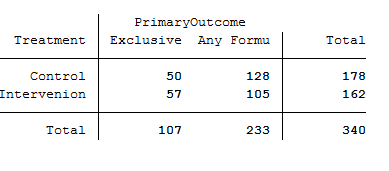
[Name of the Institution]

***Question 1 [16 marks]***

***(a) Use the tab command in Stata to create a 2x2 contingency table to display the relationship between Treatment Group and Primary\_Outcome\_6months and provide your output. [2 marks]***

***Solution***

Following is a table which is showing relationship between Treatment group and Primary\_ Outcome \_ 6 months

****

**Figure 1.:** 2\*2 Contingency table showing the relationship between

treatment group and Primary outcome \_6 months

***(b) State the Null and Alternative Hypotheses for the primary research question of interest in this study. [Hint: The lectures will be more helpful than the publication here.] [2 marks, 2-3 sentences]***

***Solution***

*Null hypothesis*

Complex intervention on knowledge, skills and social support will not increase exclusive breastfeeding duration among Lebanese women.

**H0**: There is no difference in the exclusive breast feeding among control and intervention *group.*

*Alternate Hypothesis*

Complex intervention on knowledge, skills and social support will increase exclusive breastfeeding duration among Lebanese women.

**Ha**: There is a difference in exclusive breast feeding in control and intervention group.

**c*) Calculate and report the expected values in the contingency table (to 2 decimal places), if the Null Hypothesis is true and show your working. [2 marks]***

***Solution***

Expected values are calculated by row total\*column total)/n, where n = Total observations included in the above table.

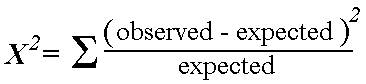
|  |  |  |  |
| --- | --- | --- | --- |
|  | **6 Months BBQ total score in categories** | | **Total** |
| **Treatment group** | **Positive** | **Negative** |  |
| **Control** | 56.02 | 121.98 | 178 |
| **Intervention** | 50.98 | 111.02 | 162 |
| **Total** | 107 | 233 | 340 |

**Figure 2**: Contingency table showing expected valued if null hypothesis is true.

***(d) We can use the information from parts (a), (b) and (c) above to calculate the 𝜒 2 statistic for this table to show that it is 1.98. Explain the steps you would take to manually calculate this statistic by hand or using Excel. [You don’t have to actually do the calculations; just state how you would perform them]. [3 marks]***

***Solution***

By using the following formula, the 𝜒 2 statistic will be calculated



*X²* = (50-56.02) ²/56.02 + (128-121.98) ²/121.98+ (57-50.98) ²/50.98+ (105-111.02) ²/111.02 = 1.98

Following are a steps that will be taken to calculate the chi square statistics.

*Step 1:*

Observed statistics will be calculated

*Step 2:*

Expected values will be generated in 2 by 2 contingency table

*Step 3*

These values will be inserted in to the equation shown above and X² value will be calculated

***(e) What is the number of degrees of freedom for this test statistic? [1 mark]***

***Solution***

*Formula to calculate the degree of freedom*

Degree of freedom is calculated by (number of columns minus one) x (number of rows minus one)

For the above table degree of freedom is equal to

(2-1) \*(2-1) = 1.

***(f) Use the tabi command in Stata to reproduce the chi-square test you carried out in Question 1. Check that your results are consistent. Include your output and note (using highlighting, arrows or similar) the parts of the output that identify the chi-square test statistic and the P value. [Hint: The P value should be consistent with that reported in the paper.] [4 marks]***

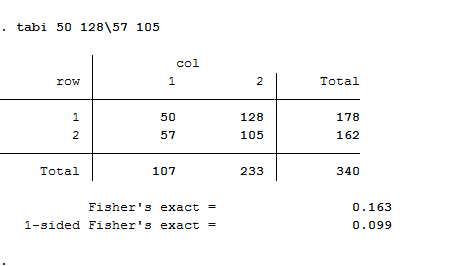
***Solution***

In the Stata Tabi command is used to reproduce the chi square test result which was generated in part 1.

*Command given in Stata*

***tabi 50 128/57 105***

*Result*



***(g) Based on your result from (f), comment on the strength of the evidence against the Null Hypothesis and what can subsequently be concluded about the primary research question. [2 marks, 2-3 sentences]***

***Solution***

Based on the result it is concluded that there is a significant difference in the treatment and control group to exclusive breast feeding. Null hypothesis is rejected as those who were in the intervention group were practicing exclusive feeding more as compared to the women in control group.

**Question 2 [22 marks]**

**(a) One of the first pieces of information that readers typically see in a trial publication is a table of baseline characteristics of the participants by treatment group. In Nabulsi M et al. this is labelled Table 1. In the publication authors comment that while the majority of characteristics in this table are balanced, there are a few that are not. Briefly explain why this is possible despite the fact that participants were randomly allocated to treatment groups. [2 marks, 2-3 sentences]**

***Solution***

**Answer**

The randomization process is basically carried out to balance assignment between the different groups, in order to avoid systematic error. However, it is also recognized that groups that are assigned through the process of randomization are not same completely. The few characteristics that are not balance are due to the chance bias.

**(b) The mean and standard deviation of the duration of previous breastfeeding (in months) for women allocated to each treatment group (taken from Table 1 in Nabulsi M et al.) can be seen below: Treatment Group Mean (months) Standard Deviation (months) Number of participants Control 11.86 8.25 188 Intervention 7.82 6.02 174 [Note: Women with no previous baby or who did not breastfeed a previous baby will necessarily have a zero duration recorded.] While statistical tests for differences of baseline characteristics should be avoided in practice, they are still commonly reported in publications and for this assignment, we shall assume that this hypothesis test is necessary. State an appropriate Null Hypothesis and the Alternative Hypothesis to look for evidence of a significant difference in average length of previous breastfeeding for Control minus Intervention. [2 marks, 2-3 sentences]**

**Solution**

*Null hypothesis*

 There is no difference between the duration of previous breastfeeding (in months) for women in control and treatment group with respect to means

H0: μ1-μ2= 0

*Alternate hypothesis*

There is a difference between the duration of previous breastfeeding (in months) for women in control and treatment group with respect to means

**H1: μ1 ≠ μ2**

**(c) Calculate and state the difference in sample mean duration of previous breastfeeding in months (Control – Intervention)? [1 mark]**

**Solution**

Difference in sample mean duration = 11.86- 7.82= 4.04

They are not close there is an evidence of difference in sample mean duration of previous breast feeding in months (controls- Intervention)

**(d) State the expected difference in sample mean duration of previous breastfeeding in months (Control – Intervention) under the Null Hypothesis. [1 mark]**

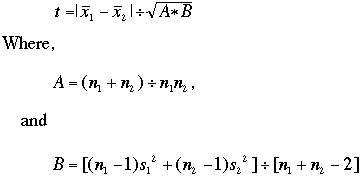
**Solution**

Under null hypothesis as μ1-μ2= 0  therefore the expected difference in sample mean duration of previous breast feeding in months (Control – intervention) will also be zero.

**(e) Using the equations seen in the lectures, the estimated standard error of the difference in sample means (i.e. 𝑒𝑠𝑡𝑖𝑚𝑎𝑡𝑒𝑑[𝑠.𝑒. (𝑥̅𝐶𝑜𝑛𝑡𝑟𝑜𝑙 − 𝑥̅𝐼𝑛𝑡𝑒𝑟𝑣𝑒𝑛𝑡𝑖𝑜𝑛)]) can be calculated by hand to be 0.76. Using this information, follow the steps below to conduct an independent samples Student’s t-test to assess the evidence of a significant difference in average length of previous breastfeeding for Control minus Intervention. i. Calculate the value of the test statistic for the difference in mean length of previous breastfeeding. Show your working (it may be hand-written, typed or a screen shot from Excel)! [3 marks]**

**Solution**

Following formula is used to calculate the independent samples Student’s t-test



A= (n1+n2)/n1n2

A= (188+174)/188\*174= 0.011

B= [(188-1) (8.25) (8.25) + (174-1) (6.02) (6.02)]/ (188+174)-2 =52.7

t= 11.86-7.82/sqrt (0.011\*52.7) = 6.96

**ii. State the number of degrees of freedom for this test statistic. [1 mark]**

**Solution**

Degree of freedom for this test statistic is calculated as follows

Df= (n1+n2) -2

= (188+174)-2 = 360

**iii. Draw a rough sketch of the distribution which your test statistic would follow if the Null Hypothesis were true, and shade the area corresponding to the probability of observing values of the test statistic as or more extreme than the observed test statistic. [2 marks]**

**Solution**

Following is a rough sketch of the distribution which the test statistic follows when null hypothesis is true**.**

Retain the Null hypothesis

Rejection region: alpha= 0.025

Rejection region: alpha= 0.025

-2 -1 0 +1 +2

**iv. Use the appropriate command in Stata to find the probability of observing a test statistic as or more extreme than the value from Question 2 (e) part i, if the Null Hypothesis were true. Display the command and its output. [1 mark**

**Solution**

**v. Use the ttesti command in Stata to repeat the comparison of mean duration of previous breastfeeding between the Control and Intervention groups. Display the command and output and note (using highlighting, arrows or similar) the parts of the output that identify values stated or calculated by you in previous parts of Question 2 (e). [4 marks]**

**Solution**

**vi. Based on your results from (iii and iv), comment on the strength of the evidence against the Null Hypothesis. [2 marks, 2-3 sentences]**

**Solution**

Based on the results, the value of t came out to be 6.96. There are (188+174-2] degree of freedom. Based on the results null hypothesis is rejected and there is a significant association present between the women with previous breast feeding experience and exclusive breast feeding.

**vii. Given what you have learned about mean duration of previous breastfeeding, comment on the potential impact this may have on the comparison of the primary outcome. [3 marks, 3-4 sentences]**

**Solution**

History of a longer duration of previous breastfeeding is a positive predictors of EBF The women with previous breast feeding experience will practice the exclusive breast feeding more. Less support is required to promote the exclusive breast feeding among them. They will more likely to practice the exclusive breast practice as compare to women who have not previous experience of breast feeding.