**Come back at 12.55pm**

**GER1000 QUANTITATIVE REASONING**

**TUTORIAL 4**

*Please work on the problems before coming to class. In class, you will engage in group work.*

**Question 1 – Group 4**

Suppose a country has 2.9 million adults, comprising of young adults (age 21 to 65) and old adults (more than 65 years old). The table classifies both groups by whether a person is an alcoholic. For example, there are 600,000 young adults who are alcoholic.

|  |  |  |
| --- | --- | --- |
|  | Alcoholic (thousand) | Non-alcoholic (thousand) |
| Young adults | 600 | 1800 |
| Old adults | 50 | 450 |

This is a census.

1. An investigator writes: “The odds for alcoholism among young adults is estimated as 3 times the odds for alcoholism among old adults.” Explain how the number is obtained. Comment on the word “estimated” in the sentence.

Odds(alcoholic|young) = 600000/180000=1/3

Odds(alcoholic|old) = 50000/450000 = 1/9

OR = 3 Population’s parameter. This is not an estimate, this is the actual population’s OR.

Assuming that the values given in the table are rounded off, “estimated” is used because the precise number of alcoholics is rounded off to the nearest thousand. So, the actual odds ratio cannot be determined with full certainty, without the actual number.

However, assuming that the population is accurate as provided: The use of the word ”estimated” would be inappropriate. This is because this study used the entire population instead of using a sample from the population.

1. Investigators of a study randomly choose 10% of the alcoholics, and 20% of the non-alcoholics, from the country. The chosen individuals are classified as young or old adults through interviews. Is the risk ratio for alcoholism between young adults and old adults in the sample smaller than, or larger than, the population risk ratio? Support your answer with some calculations.

Risk ratio of population = (600000/2400000)/(50000/500000) = **2.5**

This is a case control study

Sample’s table

|  |  |  |
| --- | --- | --- |
|  | Alcoholic (thousand) | Non-alcoholic (thousand) |
| Young adults | 60 | 360 |
| Old adults | 5 | 90 |

Sample alcoholics: 60000 young; 5000 old

Sample non-alcoholic: 360000 young; 90000 old

Risk ratio of sample: (60000/420000)/(5000/95000) = **2.71**

Thus, Risk ratio is larger in sample than that of the population. The population’s RR is overestimated.

**Question 2 - Group 2**

A game about the impact of climate change on human life uses ten independent rolls of a fair die to simulate the weather for ten years. If the die shows one spot or six spots, then there is a crisis.

Dice 1,2,3,4,5,6

P(crisis) = 2/6

P(no crisis) = 4/6 = 2/3

1. Calculate the probability that there is no crisis in a ten-year period, to 2 significant digits.

P(no crisis in 10 years) = (2/3)^10 = 0.017 (2.s.f) => +10

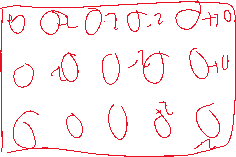
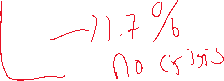
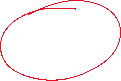
P(at least one crisis) = 1 – 0.017 = 0.983 => -2

1. The quality of life increases by 10 points if there is no crisis in a ten-year period, and decreases by 2 points otherwise. Calculate the average amount by which the quality of life increases after ten years. Interpret this number in terms of a large number of worlds independently controlled by this game.

V(A)xP(A) + V(B)xP(B)

10x0.017 + -2x(1-0.017) = -1.80

Over a large number of worlds, the average quality of life is –1.80 points.



**Question 3 - Group 5**

A taxi was involved in a hit-and-run accident at night. Two taxi companies, Green and Blue, operate in the city. An eye-witness reported the taxi involved as Blue. You are given the following data:

* There are a total of 10,000 taxis in the city. 85% of them are Green, 15% are Blue.
* When presented with a Blue taxi at night, the eye-witness reports it as Blue 8 out of 10 times; when presented with a Green taxi at night, the eye-witness reports it as Blue 3 out of 10 times.

What is the likelihood that the taxi involved in the hit-and-run is Blue, given the eye-witness’ report? Does it seem too low? How does it compare with (i) the overall percentage of Blue taxis in the city, (ii) the likelihood of the eye-witness reporting Blue, when presented with a Blue taxi at night?

Population characteristic

P(Blue) = 15%

Witness characteristic

Sensitivity => P(Reported Blue|Blue) = 80%

Specificity => P(Reported Green|Green) = 70%

|  |  |  |  |
| --- | --- | --- | --- |
|  | Reported Blue | Reported Green | Row totals |
| Blue | 1 200 | 300 | 1 500 |
| Green | 2 550 | 5 950 | 8 500 |
| Column totals | 3 750 | 6 250 | 10 000 |

P(Blue|Reported Blue) = 1200/ 3750 = **32%**

1. P(Blue) = 15% vs P(Blue|Reported Blue) = 32%

Using a witness increases the likelihood that the taxi involved in hit and run is blue.

(ii) P(Blue|Reported Blue) = 32% vs Sensitivity => P(Reported Blue|Blue) = 80%

The P(Blue|Reported Blue) is not only dependent on **sensitivity**. It is also dependent on the **specificity** quality of the witness and the proportion of blue taxis in the city.

Likelihood of hit and run is blue is lower than overall percentage of blue taxis in the city and a lot lower than likelihood of eyewitness reporting blue given a blue taxi.

**Question 4**

On Alan’s first visit to Macau, he decided to try his luck in at the Venetian Casino Resort. He played a game that is supposed to offer a 40% chance of winning. Out of five independent plays, he won the first 4 games and lost the last game. His friend Brad suspects that the game is rigged in Alan’s favour and decides to carry out a hypothesis test.

1. State Brad’s null and alternative hypotheses. - Group 1

Null: game is not rigged in Alan’s favor and he has 40% chance of winning

Alternative: game is rigged in Alan’s favour

1. If the null hypothesis is true, what is the chance of winning 4 games in a row and losing the next one? - Group 3

P(wwwwl) = (0.4)^4 x 0.6 = 0.01536

1. Calculate the P value based on Alan’s data, to one significant figure. - Group 1

4w 1 l

Wwwwl

Wwwlw

Wwlww

Wlwww

Lwwww

More extreme: wwwww => (0.4)^5

P-value = 0.01536\*5 + (0.4)^5 = 0.09

1. What can Brad conclude, based on 5% level of statistical significance? - Group 3

We have insufficient evidence to reject the null hypothesis.

There is no conclusion, game may or may not be rigged.

Conclusion: no conclusion, the game may or may not be rigged. We can still assume that the game is not rigged