

fishers_odds_ratio

January 31, 2023

```
[1]: import numpy as np
      from scipy.stats import fisher_exact, norm

[2]: def calc_odds_ratio(a, b, c, d):
      odds_ratio = (a * d) / (b * c)
      p_value = fisher_exact([[a, b], [c, d]])[1]
      se = np.sqrt(1/a + 1/b + 1/c + 1/d)
      z = np.log(odds_ratio) / se
      confint = np.exp(norm.ppf(0.025) * se), np.exp(norm.ppf(0.975) * se)
      return odds_ratio, p_value, confint

[3]: #Age
      odds_ratio, p_value, confint = calc_odds_ratio(87, 32, 98, 263)

      print("The Fisher's Exact Associatio for age =")
      print("Odds ratio: ", "{:.3f}".format(odds_ratio))
      print("P-value: ", "{:.3e}".format(p_value))
      print("confint_lower", "{:.3f}".format(confint[0]))
      print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for age =
Odds ratio:  7.296
P-value:    1.179e-18
confint_lower 0.627
confint_upper 1.595
```

The Fisher's exact odds ratio of 7.296 indicates that the odds of being a positive case in the group with less than 1 year is 7 times higher compared to the group with greater than 1 year.

The p-value of 1.179e-18 is extremely small, which suggests strong evidence against the null hypothesis. In other words, the association between the two groups is statistically significant.

The confidence interval provides a range of values that is likely to contain the true proportion of positive cases with a certain level of confidence, in this case 95% (as defined by the alpha value of 0.05). The lower and upper bounds of the confidence interval are 0.627 and 1.595 respectively, indicating that the true proportion of positive cases in the two groups are likely to be between these values.

```
[4]: #Sex
odds_ratio, p_value, confint = calc_odds_ratio(25, 94, 95, 266)

print("The Fisher's Exact Associatio for sex =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for sex =
Odds ratio:  0.745
P-value:    2.733e-01
confint_lower 0.607
confint_upper 1.648
```

The odds ratio of 0.745 indicates that the odds of a positive outcome are 0.745 times higher in the female group compared to the male group. This means that for every one unit increase in the odds of a positive outcome in the male group, the odds of a positive outcome increase by 0.745 units in the female group.

The p-value of 0.273 suggests that the results are not statistically significant at the commonly used significance level of 0.05. This means that there is a 27.3% chance of observing the results by chance if there is no real difference between the male and female groups.

The confidence interval of (0.607, 1.648) shows that with 95% confidence, the true odds ratio lies between 0.607 and 1.648. This means that if the study were repeated many times with different samples, 95% of the confidence intervals would contain the true odds ratio. The lower end of the interval (0.607) suggests that the odds of a positive outcome are 0.607 times higher in the female group compared to the male group, while the upper end of the interval (1.648) suggests that the odds of a positive outcome are 1.648 times higher in the female group compared to the male group.

```
[5]: #presence of cat
odds_ratio, p_value, confint = calc_odds_ratio(106, 13, 56, 305)

print("The Fisher's Exact Associatio for Presence of Cat =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for Presence of Cat =
Odds ratio:  44.409
P-value:    2.802e-49
confint_lower 0.526
confint_upper 1.901
```

The odds ratio of 44.409 indicates that the odds of a positive outcome are 44.409 times higher in the group with the presence of cats compared to the group without the presence of cats. This means that for every one unit increase in the odds of a positive outcome in the group without the presence of cats, the odds of a positive outcome increase by 44.409 units in the group with the

presence of cats.

The p-value of 2.802e-49 suggests that the results are highly statistically significant at the commonly used significance level of 0.05. This means that there is a very small chance of observing the results by chance if there is no real difference between the groups with and without the presence of cats.

The confidence interval of (0.526, 1.901) shows that with 95% confidence, the true odds ratio lies between 0.526 and 1.901. This means that if the study were repeated many times with different samples, 95% of the confidence intervals would contain the true odds ratio. The lower end of the interval (0.526) suggests that the odds of a positive outcome are 0.526 times higher in the group with the presence of cats compared to the group without the presence of cats, while the upper end of the interval (1.901) suggests that the odds of a positive outcome are 1.901 times higher in the group with the presence of cats compared to the group without the presence of cats.

```
[6]: #cat in contact with cattle
odds_ratio, p_value, confint = calc_odds_ratio(114, 5, 30, 331)

print("The Fisher's Exact Associatio for Cat in Contact with Cattle =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for Cat in Contact with Cattle =
Odds ratio: 251.560
P-value: 1.219e-74
confint_lower 0.379
confint_upper 2.639
```

The odds ratio of 251.560 suggests that cattle that have contact with cats have 251.560 times higher odds of being infected with *T. gondii* compared to cattle without contact with cats. The p-value of 1.219e-74 is very small, indicating strong evidence against the null hypothesis (that there is no association between contact with cats and *T. gondii* infection in cattle). The confidence interval (0.379, 2.639) gives the range of values for the true odds ratio with a 95% level of confidence, meaning that if we repeated the study many times, 95% of the intervals would contain the true odds ratio.

```
[7]: #Cat with contact in drinking water
odds_ratio, p_value, confint = calc_odds_ratio(87, 32, 34, 327)

print("The Fisher's Exact Associatio for Cat with Contact in Drinking Water =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for Cat with Contact in Drinking Water =
Odds ratio: 26.148
P-value: 3.055e-40
confint_lower 0.584
```

```
confint_upper 1.712
```

The results indicate that there is a significant association between *T. gondii* infection among cattle and drinking water with cats. The odds of *T. gondii* infection among cattle drinking water with cats is 26.148 times higher compared to cattle not drinking water with cats. The p-value of 3.055e-40 indicates that the association is statistically significant. The confidence interval (0.584, 1.712) suggests that the true odds ratio is likely to lie between 0.584 and 1.712, with 95% confidence.

```
[8]: #Presence of rats
odds_ratio, p_value, confint = calc_odds_ratio(112, 7, 242, 119)

print("The Fisher's Exact Associatio for Presence of Rats =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for Presence of Rats =
Odds ratio:  7.868
P-value:    3.008e-10
confint_lower 0.452
confint_upper 2.213
```

The odds ratio of 7.868 suggests that cattle with the presence of rats on the farm are approximately 8 times more likely to be infected with *T. gondii* compared to cattle without the presence of rats on the farm.

The p-value of 3.008e-10 is less than 0.05, indicating that the association between the presence of rats and *T. gondii* infection in cattle is statistically significant.

The confidence interval of 0.452 to 2.213 shows that we can be confident (with 95% confidence) that the true odds ratio lies between these values.

```
[9]: #House type
odds_ratio, p_value, confint = calc_odds_ratio(107, 12, 283, 78)

print("The Fisher's Exact Associatio for House Type =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for House Type =
Odds ratio:  2.458
P-value:    4.405e-03
confint_lower 0.524
confint_upper 1.910
```

The results show that cattle living in a pen house are more likely to be infected with *T. gondii* compared to cattle living in a fence house. The odds ratio of 2.458 indicates that the odds of *T. gondii* infection are approximately 2.5 times higher for cattle living in a pen house compared to

those living in a fence house. The statistical significance of this association is indicated by the p-value of 4.405e-03, which is less than 0.05. The 95% confidence interval of 0.524 to 1.910 provides a range within which we can be confident that the true odds ratio lies.

```
[10]: #Water source
odds_ratio, p_value, confint = calc_odds_ratio(65, 54, 241, 120)

print("The Fisher's Exact Associatio for Water Source =")
print("Odds ratio: ", "{:.3f}".format(odds_ratio))
print("P-value: ", "{:.3e}".format(p_value))
print("confint_lower", "{:.3f}".format(confint[0]))
print("confint_upper", "{:.3f}".format(confint[1]))
```

```
The Fisher's Exact Associatio for Water Source =
Odds ratio:  0.599
P-value:    2.079e-02
confint_lower 0.656
confint_upper 1.525
```

The odds ratio of 0.599 suggests that cattle drinking water from a well are approximately 60% less likely to be infected with *T. gondii* compared to cattle drinking water from a pond.

The p-value of 2.079e-02 is less than 0.05, indicating that the association between drinking water source and *T. gondii* infection in cattle is statistically significant.

The confidence interval of 0.656 to 1.525 shows that we can be confident (with 95% confidence) that the true odds ratio lies between these values.