WILD 5200/6200; FORY 7200

Sep 11, 2024

Worksheet 2 & 3: Measures of disease frequency

Type the questions and your answers in the quarto document you have created. You can insert the table in your Quarto document. Available in the audiseco repository on Github. Link: https://github.com/anyadoc/audiseco/blob/main/ss_table.png

Please upload your quarto document to your Github repository after completing this worksheet.

1) Find the sample size necessary to detect a disease at or above 5% prevalence in a herd of 5000 deer. The desired level of confidence is 95%.

Sample Size necessary: 59 deer

2) Using the hypergeometric distribution, compare the sample size requirements for the following scenarios:

a) 95% confidence in detecting a disease at 1% prevalence in a large population

- **In a large population (Infinite), a sample size of 299 would be needed for detecting disease at 1% prevalence and 95% Confidence.**

b) 99% confidence in detecting a disease at 2% prevalence in a large population

- **In a large population (Infinte), a sample size of 229 would be needed to detect disease at a 2% prevalence and 99% Confidence level.**

3) You sample 300 deer from a high fenced facility that has a herd of 4000. None of the 300 have antibodies against Epizootic Hemorrhagic Disease. Based on this finding, you are 95% confident that the prevalence of exposure to EHD in this facility is __<___ (<, =, >) __1__%.

- Population size: 4000
- Sample size: 300
- Confidence level: 95%
- Number of positive cases detected: 0
- **95% confident that the prevalence of exposure to EHD in this facility is <1% because a sample size of 288 has a 1% prevalence**

WILD 5200/6200; FORY 7200

Sep 11, 2024

Worksheet 2 & 3: Measures of disease frequency

4) An outbreak of 'Lame disease' has been reported in wild turkey populations of Alabama. Turkeys from three counties were sampled for lameness (Lee, Chambers and Russell counties).

a) Analyze the sample data to determine prevalence of lameness in each county. Provide confidence interval for prevalence using Sterne's method.

- **Chambers: Prevalence = 8/30 = 0.267 or 26.7% at a 95% Confidence Interval**

```
    Two-sided p-value:  0.1987
_____________________________________________________________________________

 QPweb                Sun Sep 15 21:27:16 2024

    Exact confidence limits for the prevalence (Sterne's method, new algorithm)
    Confidence level = 95%
                                N  Infected  Prev.     Lower CL  Upper CL
DeerLameness_Chamb$lameness    30       8   0.267        0.131     0.449
```

-
- **Lee: Prevalence = 30/150 = 0.200 or 20% at a 95% Confidence Interval**

```
    Exact confidence limits for the prevalence (Sterne's method, new algorithm)
    Confidence level = 95%
                                  N  Infected  Prev.     Lower CL  Upper CL
DeerLameness_Lee$lameness       150      30   0.200        0.143     0.273
```

-
- **Russel: Prevalence = 3/96 = 0.031 or 3.1 % at a 95% Confidence Interval**

```
    Exact confidence limits for the prevalence (Sterne's method, new algorithm)
    Confidence level = 95%
                                   N  Infected  Prev.     Lower CL  Upper CL
DeerLameness_Russe$lameness       96       3   0.031        0.009     0.087
```

-

b) Compare prevalences using Fisher's exact test.

- **Fisher's Exact Test Results:**

```
    Comparison of prevalences by Fisher's exact test

                                   N  Infected Non-inf.   Prev.%
DeerLameness_Chamb$lameness       30         8       22     26.7
DeerLameness_Lee$lameness        150        30      120     20.0
DeerLameness_Russe$lameness       96         3       93      3.1

    Two-sided p-value: <0.0001
```

-
- **Confirms (a) → Each Prevalence**

The sampling data is available in the audiseco repository on Github. You can import the data in your R Notebook using the links provided here:

WILD 5200/6200; FORY 7200

Sep 11, 2024

Worksheet 2 & 3: Measures of disease frequency
Lee County:
https://raw.githubusercontent.com/anyadoc/audiseco/main/DeerLameness_Lee.csv

Chambers County:
https://raw.githubusercontent.com/anyadoc/audiseco/main/DeerLameness_Chambers.csv

Russell County:
https://raw.githubusercontent.com/anyadoc/audiseco/main/DeerLameness_Russell.csv

Use Quantitative Parasitology (QPweb) for calculating confidence intervals and comparing prevalences: http://www.zoologia.hu/qp/qp.html

Step-by-step instructions for QPweb are available here:
https://github.com/anyadoc/audiseco/blob/main/Using%20QPweb.pdf

5) Jane was studying the spread of Pink Spot Disease (PSD) in a pond that had 60 fish. She introduced an infectious fish in the pond on day 0. Her observations are as follows:

Day 1: 3 fish infected

Day 2: 4 fish infected

Day 3: 7 fish infected

Day 4: 11 fish infected

Day 5: 20 fish infected

Note that there is no disease-related mortality. The observations denote all infected fish in the pond (including previously infected fish).

a) Find incidence on day 1, day 2 and day 5. Write the unit for incidence.

- Prev**ALL**ence: **ALL** cases/population total
- I**N**cidence: **New** cases/population at risk

Day 1:

- New cases: 3
- Population at risk: 60 (there were none infected on day 1)

Worksheet 2 & 3: Measures of disease frequency

- IN**c**idence: **New** cases/population at risk → **3/60 = 0.05 x 100 = 5 cases per 100 fish per day**

Day 2:

- New cases: (4 total - 3 already infected) = 1
- Population at risk: 60 (there were none infected on day 1) - 3 (already infected) = 57
- IN**c**idence: **New** cases/population at risk → **1/57 = 0.0175 x 100 = 1.75 cases per 100 fish per day**

Day 5:

- New cases: (20 total - 11 already infected) = 9
- Population at risk: 60 (there were none infected on day 1) - 11 infected = 49 at risk
- IN**c**idence: **New** cases/population at risk → **9/49 = 0.1837 x 100 = 18.37 cases per 100 fish per day**

b) What is the prevalence of PSD on day 3 and day 5?

- Prev**ALL**ence: **ALL** cases/population total

Day 3:

- All cases: 7
- Population: 60
- **7/60 = 0.1167 x 100 = 11.67 % prevalence**

Day 5:

- All cases: 20
- Population: 60
- **20/60 = 0.3333 x 100 = 33.33 % prevalence**

c) On which day is the prevalence the same as the incidence (except the units)?

- **On day one, the prevalence and incidence are the same**
- Prevalence: 3/60 = 0.05
- Incidence: 3/60 = 0.05