

DATA 605 Assignment 5

Warner Alexis

2024-02-24

Problems

1. (Bayesian). A new test for multinucleoside-resistant (MNR) human immunodeficiency virus type 1

(HIV-1) variants was recently developed. The test maintains 96% sensitivity, meaning that, for those with the disease, it will correctly report “positive” for 96% of them. The test is also 98% specific, meaning that, for those without the disease, 98% will be correctly reported as “negative.” MNR HIV-1 is considered to be rare (albeit emerging), with about a .1% or .001 prevalence rate. Given the prevalence rate, sensitivity, and specificity estimates, what is the probability that an individual who is reported as positive by the new test actually has the disease? If the median cost (consider this the best point estimate) is about \$100,000 per positive case total and the test itself costs \$1000 per administration, what is the total first-year cost for treating 100,000 individuals?

Solution

We are going to use the Bayes’ Theorem which defines as :

$$P(A|B) = \frac{P(A|B)P(A)}{P(B)}$$

$$P(Disease|Positive) = \frac{P(Positive|Disease)P(Disease)}{P(Positive)}$$

$$P(Disease) = 0.001$$

$$P(Negative|NoDisease) = 0.98$$

$$P(Positive) = P(Positive|Disease) * P(Disease) + P(Positive|NoDisease) * P(NoDisease)$$

$$P(Positive) = (0.96 * 0.001) + (0.02 * 0.999)$$

$$P(Positive) = 0.02094$$

$$P(Disease|Positive) = \frac{0.96 * 0.001}{0.02094}$$

$$P(Disease|Positive) = 0.0458$$

To treat 100,000 Individuals we need: $0.0458 * 100000 = 4580$ cases

Total cost for treating 100,000 Individuals is $4580 * 1000 = \$ 4,580,000$

```
library(LaplacesDemon)
PrA <- c(0.96)

PrBA <- c(0.001, 0.02094)

BayesTheorem(PrA, PrBA)
```

```
## [1] 0.04557885 0.95442115
## attr(,"class")
## [1] "bayestheorem"
```

2. (Binomial). The probability of your organization receiving a Joint Commission inspection in any given

month is .05. What is the probability that, after 24 months, you received exactly 2 inspections? What is the probability that, after 24 months, you received 2 or more inspections? What is the probability that your received fewer than 2 inspections? What is the expected number of inspections you should have received? What is the standard deviation?

Solution

Probability of receiving a Joint Commission inspection in any given month $p = 0.05$ Number of trials in months $n = 24$

Probability of exactly 2 inspections Given yjay Probability mass function is : $P(X = k) = \binom{n}{k} * p^k * (1 - p)^{n - k}$

$$P(X = k) = \binom{24}{2} * 0.05^2 * (1 - 0.05)^{24 - 2}$$

$$P(X = k) = 0.2232381$$

```
# Given data
p <- 0.05 # Probability of receiving an inspection in any given month
n <- 24   # Number of months

# 1. Probability of exactly 2 inspections
prob_2_inspections <- dbinom(2, size = n, prob = p)
cat("1. Probability of exactly 2 inspections:", prob_2_inspections, "\n")
```

```
## 1. Probability of exactly 2 inspections: 0.2232381
```

Probabilities of 2 more inspections or $P(X > 2) = 1 - P(X < 2)$

```
# 2. Probability of 2 or more inspections
prob_2_or_more_inspections <- 1 - pbinom(1, size = n, prob = p)

cat("2. Probability of 2 or more inspections:", prob_2_or_more_inspections, "\n")
```

```
## 2. Probability of 2 or more inspections: 0.3391827
```

Probabilities of fewer than 2 inspections

```
# 3. Probability of fewer than 2 inspections
prob_less_than_2_inspections <- pbinom(1, size = n, prob = p)
cat("3. Probability of fewer than 2 inspections:", prob_less_than_2_inspections, "\n")
```

```
## 3. Probability of fewer than 2 inspections: 0.6608173
```

Probabilities of the number of inspections

```
# 4. Expected number of inspections
expected_value = n * p

# 5. Standard deviation

# 5. Standard deviation
std_deviation <- sqrt(n * p * (1 - p))
cat("4. Expected number of inspections:", expected_value, "\n")
```

```
## 4. Expected number of inspections: 1.2
```

```
cat("5. Standard deviation:", std_deviation, "\n")
```

```
## 5. Standard deviation: 1.067708
```

3. (Poisson). You are modeling the family practice clinic and notice that patients arrive at a rate of 10 per

hour. What is the probability that exactly 3 arrive in one hour? What is the probability that more than 10 arrive in one hour? How many would you expect to arrive in 8 hours? What is the standard deviation of the appropriate probability distribution? If there are three family practice providers that can see 24 templated patients each day, what is the percent utilization and what are your recommendations?

Probability that exactly 3 arrive in one hour:

$$P(X) = \frac{\lambda^x}{x!} e^{-\lambda}$$

```
# Given data
lambda <- 10 # Average rate of patient arrivals per hour
hours <- 8   # Number of hours

# 1. Probability that exactly 3 arrive in one hour
prob_3_arrivals <- dpois(3, lambda)
cat("1. Probability that exactly 3 arrive in one hour:", prob_3_arrivals, "\n")
```

```
## 1. Probability that exactly 3 arrive in one hour: 0.007566655
```

Probability that more than 10 arrive in 8 hours

```
prob_more_than_10_arrivals <- 1 - ppois(10, lambda)
cat("2. Probability that more than 10 arrive in one hour:", prob_more_than_10_arrivals, "\n")
```

```
## 2. Probability that more than 10 arrive in one hour: 0.4169602
```

Expected number if arrivals in 8 hours

```
expected_arrivals <- lambda * hours

cat("3. Expected number of arrivals in 8 hours:", expected_arrivals, "\n")
```

```
## 3. Expected number of arrivals in 8 hours: 80
```

Standard Deviation

```
std_deviation <- sqrt(lambda)
cat("4. Standard deviation:", std_deviation, "\n")
```

```
## 4. Standard deviation: 3.162278
```

5 (Hypergeometric). Your subordinate with 30 supervisors was recently accused of favoring nurses. 15 of the subordinate's workers are nurses and 15 are other than nurses. As evidence of malfeasance, the accuser stated that there were 6 company-paid trips to Disney World for which everyone was eligible. The supervisor sent 5 nurses and 1 non-nurse. If your subordinate acted innocently, what was the probability he/she would have selected five nurses for the trips? How many nurses would we have expected your subordinate to send? How many non-nurses would we have expected your subordinate to send?

The Probability of selecting five nurse for the trips

```
# Given data
number_workers <- 30      # Total number of workers
nurses <- 15              # Number of nurses
non_nurses <- 15          # Number of non-nurses
trips <- 6

nurses_sent <- 5          # Number of nurses sent on trips
non_nurses_sent <- 1

# 1. Probability of selecting five nurses for the trips
prob_5_nurses <- dhyper(nurses_sent, nurses, non_nurses, trips)
cat("1. Probability of selecting five nurses for the trips:", prob_5_nurses, "\n")
```

```
## 1. Probability of selecting five nurses for the trips: 0.07586207
```

The expected nurses and non nurses sent:

```
# 2. Expected number of nurses sent
expected_nurses_sent <- trips * (nurses / number_workers)

# 3. Expected number of non-nurses sent
expected_non_nurses_sent <- trips - expected_nurses_sent

cat("2. Expected number of nurses sent:", expected_nurses_sent, "\n")
```

```
## 2. Expected number of nurses sent: 3
```

```
cat("3. Expected number of non-nurses sent:", expected_non_nurses_sent, "\n")
```

```
## 3. Expected number of non-nurses sent: 3
```

5 (Geometric). The probability of being seriously injured in a car crash in an unspecified location is about .1% per hour. A driver is required to traverse this area for 1200 hours in the course of a year. What is the probability that the driver will be seriously injured during the course of the year? In the course of 15 months? What is the expected number of hours that a driver will drive before being seriously injured? Given that a driver has driven 1200 hours, what is the probability that he or she will be injured in the next 100 hours?

```
# Given data
p <- 0.001 # Probability of being seriously injured per hour
hours_year <- 1200 # Total number of hours in a year
hours_15_months <- 15 * 30 * 24 # Total number of hours in 15 months
hours_driven <- 1200 # Number of hours already driven
remaining_hours <- 100 #

# Prob that drivers will be seriously injured
prob_driver <- pgeom(hours_year, p)
cat("1. Probabilities that the driver will seriously injured in a car crash:", prob_driver, "\n")

## 1. Probabilities that the driver will seriously injured in a car crash: 0.6992876

cat("2. Probability of being seriously injured during the course of 15 months:", pgeom(hours_driven *
remaining_hours, p), "\n")

## 2. Probability of being seriously injured during the course of 15 months: 0.7772602

cat("3. Expected number of hours before being seriously injured", 1 / p, "\n")

## 3. Expected number of hours before being seriously injured 1000

cat("4. Probability of being injured in the next 100 hours, given that 1200 hours have already been driven",
pgeom(remaining_hours, p), "\n")

## 4. Probability of being injured in the next 100 hours, given that 1200 hours have already been driven 0.0001053992
```

6 You are working in a hospital that is running off of a primary generator which fails about once in 1000

hours. What is the probability that the generator will fail more than twice in 1000 hours? What is the expected value?

```
cat("Probability that the generator would fall twice in 1000 hours", 1 - ppois(2, 1), "\n")

## Probability that the generator would fall twice in 1000 hours 0.0803014
```

7 A surgical patient arrives for surgery precisely at a given time. Based on previous analysis (or a lack of

knowledge assumption), you know that the waiting time is uniformly distributed from 0 to 30 minutes. What is the probability that this patient will wait more than 10 minutes? If the patient has already waited 10 minutes, what is the probability that he/she will wait at least another 5 minutes prior to being seen? What is the expected waiting time?

```
# Given data
min_time <- 0 # Minimum waiting time in minutes
max_time <- 30 # Maximum waiting time in minutes

# 1. Probability that the patient will wait more than 10 minutes
prob_wait_10_min <- (max_time - 10) / max_time

# 2. Probability that the patient will wait at least another 5 minutes after already waiting 10 minutes
prob_wait_5_min <- (max_time - 15) / max_time

# 3. Expected waiting time
expected_time <- (min_time + max_time) / 2

# Output the results
cat("1. Probability that the patient will wait more than 10 minutes:", prob_wait_10_min, "\n")
```

```
## 1. Probability that the patient will wait more than 10 minutes: 0.6666667
```

```
cat("2. Probability that the patient will wait at least another 5 minutes after waiting 10 minutes:", p
```

```
## 2. Probability that the patient will wait at least another 5 minutes after waiting 10 minutes: 0.5
```

```
cat("3. Expected waiting time:", expected_time, "minutes\n")
```

```
## 3. Expected waiting time: 15 minutes
```

8 Your hospital owns an old MRI, which has a manufacturer's lifetime of about 10 years (expected value). Based on previous studies, we know that the failure of most MRIs obeys an exponential distribution.

Expected value

```
expected_lifetime <- 10
cat("1.The expected value for failure time ", 1/(1/10), "\n")
```

```
## 1.The expected value for failure time 10
```

```
cat("2.The standar deviation for failure time ", 1/(1/10), "\n")
```

```
## 2.The standar deviation for failure time 10
```

```
rate <- 1 / expected_lifetime  
cat("Probability that the MRI fails before 5 years:", pexp(5, rate), "\n")
```

```
## Probability that the MRI fails before 5 years: 0.3934693
```

Expected failure time

```
cat("Expected Time before Failure :", 1 / rate, "\n")
```

```
## Expected Time before Failure : 10
```