Khyati Naik: Data 605 - HW5

Sep 30, 2023

Contents

Bayesian

Problem 1

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?

```
# Given values
sensitivity <- 0.96  # Sensitivity of the test
specificity <- 0.98  # Specificity of the test
prevalence <- 0.001  # Prevalence rate
cost_per_positive_case <- 100000  # Cost per positive case
cost_per_test <- 1000  # Cost per test
total_individuals <- 100000  # Total number of individuals

# Calculate the probability of having the disease given a positive test result using Bayes' theorem
probability_disease_given_positive <- (sensitivity * prevalence) / ((sensitivity * prevalence) + ((1 - s

# Calculate the total first-year cost for treating 100,000 individuals
```

```r
total_cost <- (probability_disease_given_positive * total_individuals * cost_per_positive_case) + (total

# Print the results
cat("Probability that an individual who is reported as positive actually has the disease:", probability_
```

```
## Probability that an individual who is reported as positive actually has the disease: 0.04584527
```

```r
cat("Total first-year cost for treating 100,000 individuals:", total_cost, "\n")
```

```
## Total first-year cost for treating 100,000 individuals: 558452722
```

Problem 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?

```r
# Given values
probability_inspection <- 0.05  # Probability of receiving an inspection
months <- 24  # Number of months

# Probability of receiving exactly 2 inspections in 24 months
probability_exactly_2 <- dbinom(2, months, probability_inspection)

# Probability of receiving 2 or more inspections in 24 months
probability_2_or_more <- pbinom(1, months, probability_inspection, lower.tail = FALSE)

# Probability of receiving fewer than 2 inspections in 24 months
probability_fewer_than_2 <- pbinom(1, months, probability_inspection)

# Expected number of inspections
expected_inspections <- months * probability_inspection

# Standard deviation of inspections
standard_deviation <- sqrt(months * probability_inspection * (1 - probability_inspection))

# Print the results
cat("Probability of receiving exactly 2 inspections:", probability_exactly_2, "\n")
```

```
## Probability of receiving exactly 2 inspections: 0.2232381
```

```r
cat("Probability of receiving 2 or more inspections:", probability_2_or_more, "\n")
```

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

```r
cat("Probability of receiving fewer than 2 inspections:", probability_fewer_than_2, "\n")
```

```
## Probability of receiving fewer than 2 inspections: 0.6608173
```

```r
cat("Expected number of inspections:", expected_inspections, "\n")
```

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

```r
cat("Standard deviation of inspections:", standard_deviation, "\n")
```

```
## Standard deviation of inspections: 1.067708
```

Problem 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?

```r
# Given values
arrival_rate_per_hour <- 10  # Patients arrive at a rate of 10 per hour
hours <- 1  # Number of hours for the first two questions
hours_8 <- 8  # Number of hours for the third question
providers <- 3  # Number of family practice providers
patients_per_provider_per_day <- 24  # Number of patients each provider can see per day

# Probability that exactly 3 patients arrive in one hour
probability_3_patients <- dpois(3, lambda = arrival_rate_per_hour * hours)

# Probability that more than 10 patients arrive in one hour
probability_more_than_10 <- 1 - ppois(10, lambda = arrival_rate_per_hour * hours)

# Expected number of patients arriving in 8 hours
expected_patients_8_hours <- arrival_rate_per_hour * hours_8

# Standard deviation of the Poisson distribution
standard_deviation <- sqrt(arrival_rate_per_hour * hours_8)

# Total patients that can be seen by all providers in a day
total_patients_per_day <- providers * patients_per_provider_per_day

# Percent utilization
percent_utilization <- (expected_patients_8_hours / total_patients_per_day) * 100

# Print the results
cat("Probability that exactly 3 patients arrive in one hour:", probability_3_patients, "\n")
```

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

```r
cat("Probability that more than 10 patients arrive in one hour:", probability_more_than_10, "\n")
```

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

```r
cat("Expected number of patients arriving in 8 hours:", expected_patients_8_hours, "\n")
```

```
## Expected number of patients arriving in 8 hours: 80
```

```r
cat("Standard deviation of the appropriate probability distribution:", standard_deviation, "\n")
```

```
## Standard deviation of the appropriate probability distribution: 8.944272
```

```r
cat("Percent utilization:", percent_utilization, "%\n")
```

```
## Percent utilization: 111.1111 %
```

Since utilization rate is above 100%, I recommend to hire another family provider.

Problem 4

(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?

```r
# Given values
total_workers <- 30  # Total number of workers
total_nurses <- 15  # Total number of nurses
total_non_nurses <- total_workers - total_nurses  # Total number of non-nurses
trips_available <- 6  # Total number of company-paid trips
trips_taken_nurses <- 5  # Number of trips taken by nurses
trips_taken_non_nurses <- 1  # Number of trips taken by non-nurses

# Probability of selecting 5 nurses out of 6 trips taken innocently
probability_5_nurses <- dhyper(trips_taken_nurses, total_nurses, total_non_nurses, trips_available)

# Expected number of nurses selected for the trips
expected_nurses_selected <- (total_nurses * trips_available) / total_workers

# Expected number of non-nurses selected for the trips
expected_non_nurses_selected <- (total_non_nurses * trips_available) / total_workers

# Print the results
cat("Probability of selecting 5 nurses out of 6 trips taken innocently:", probability_5_nurses, "\n")
```

```
## Probability of selecting 5 nurses out of 6 trips taken innocently: 0.07586207
```

```r
cat("Expected number of nurses selected for the trips:", expected_nurses_selected, "\n")
```

```
## Expected number of nurses selected for the trips: 3
```

```r
cat("Expected number of non-nurses selected for the trips:", expected_non_nurses_selected, "\n")
```

```
## Expected number of non-nurses selected for the trips: 3
```

Problem 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?

```r
# Given values
probability_injury_per_hour <- 0.001  # Probability of being seriously injured per hour
hours_per_year <- 1200  # Hours driven in a year
hours_per_15_months <- 1200 * 15/12  # Hours driven in 15 months
hours_until_injury <- 1 / probability_injury_per_hour  # Expected hours until injury

# Probability that the driver will be seriously injured during the course of the year
probability_injury_in_year <- 1 - (1 - probability_injury_per_hour) ^ hours_per_year

# Probability that the driver will be seriously injured in the course of 15 months
probability_injury_in_15_months <- 1 - (1 - probability_injury_per_hour) ^ hours_per_15_months

# Probability that the driver will be injured in the next 100 hours given 1200 hours driven
probability_injury_in_next_100_hours <- 1 - (1 - probability_injury_per_hour) ^ 100

# Print the results
cat("Probability that the driver will be seriously injured during the course of the year:", probability_
```

```
## Probability that the driver will be seriously injured during the course of the year: 0.6989866
```

```r
cat("Probability that the driver will be seriously injured in the course of 15 months:", probability_in
```

```
## Probability that the driver will be seriously injured in the course of 15 months: 0.7770372
```

```r
cat("Expected number of hours that a driver will drive before being seriously injured:", hours_until_in
```

```
## Expected number of hours that a driver will drive before being seriously injured: 1000
```

```r
cat("Probability of being injured in the next 100 hours given 1200 hours driven:", probability_injury_in
```

```
## Probability of being injured in the next 100 hours given 1200 hours driven: 0.09520785
```

Problem 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?

```
# Given values
failure_rate_per_hour <- 1 / 1000  # Probability of generator failure per hour
hours <- 1000  # Total number of hours

# Probability that the generator will fail more than twice in 1000 hours
probability_fail_more_than_twice <- 1 - ppois(2, lambda = failure_rate_per_hour * hours)

# Expected number of failures in 1000 hours
expected_failures <- failure_rate_per_hour * hours

# Print the results
cat("Probability that the generator will fail more than twice in 1000 hours:", probability_fail_more_tha
```

```
## Probability that the generator will fail more than twice in 1000 hours: 0.0803014
```

```
cat("Expected number of failures in 1000 hours:", expected_failures, "\n")
```

```
## Expected number of failures in 1000 hours: 1
```

Problem 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 values
min_waiting_time <- 0  # Minimum waiting time in minutes
max_waiting_time <- 30  # Maximum waiting time in minutes
time_interval_10_min <- 10  # Waiting time interval for the first question in minutes
time_interval_5_min <- 5  # Waiting time interval for the second question in minutes

# Probability that patient will wait more than 10 minutes
probability_wait_more_than_10 <- 1 - punif(time_interval_10_min, min = min_waiting_time, max = max_waiti

# Probability that patient will wait at least another 5 minutes after waiting for 10 minutes
probability_wait_at_least_5_after_10 <- 1 - punif(time_interval_10_min + time_interval_5_min, min = min_

# Expected waiting time (mean of uniform distribution)
expected_waiting_time <- (min_waiting_time + max_waiting_time) / 2

# Print the results
cat("Probability that patient will wait more than 10 minutes:", probability_wait_more_than_10, "\n")
```

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

```r
cat("Probability that patient will wait at least another 5 minutes after waiting for 10 minutes:", proba
```

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

```r
cat("Expected waiting time:", expected_waiting_time, "minutes\n")
```

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

Problem 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. What is the expected failure time? What is the standard deviation? What is the probability that your MRI will fail after 8 years? Now assume that you have owned the machine for 8 years. Given that you already owned the machine 8 years, what is the probability that it will fail in the next two years?

```r
# Given values
lifetime_mean <- 10  # Expected lifetime in years

# Expected failure time (mean) for an exponential distribution
expected_failure_time <- lifetime_mean

# Standard deviation for an exponential distribution
standard_deviation <- expected_failure_time

# Probability that the MRI will fail after 8 years
probability_fail_after_8_years <- pexp(8, rate = 1 / expected_failure_time, lower.tail = FALSE)

# Probability that the MRI will fail in the next two years given 8 years have passed
probability_fail_in_next_2_years <- pexp(2, rate = 1 / expected_failure_time)

# Print the results
cat("Expected failure time (mean):", expected_failure_time, "years\n")
```

```
## Expected failure time (mean): 10 years
```

```r
cat("Standard deviation:", standard_deviation, "years\n")
```

```
## Standard deviation: 10 years
```

```r
cat("Probability that the MRI will fail after 8 years:", probability_fail_after_8_years, "\n")
```

```
## Probability that the MRI will fail after 8 years: 0.449329
```

```r
cat("Probability that the MRI will fail in the next two years (given 8 years have passed):", probability
```

```
## Probability that the MRI will fail in the next two years (given 8 years have passed): 0.1812692
```