Data 605 - Assignment 5

Leticia Salazar

2/24/2022

Contents

Assignment 5

Problem 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:

Key: A_1 = Actual Positive

A_2 = Actual Negative

B = Positive Test

P(A_1) = 0.001

P(B | A_1) = 0.96

P(B | A_2) = 0.02 (1 - 0.98)

P(A_2) = 0.999

```r
# 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?

# With the Bayes method try to find P(A1 | B):
paste("The probability that an individual who is reported as positive by the new test actually has the
    round(0.96 * 0.001/(0.96 * 0.001 + 0.02 * 0.999), 4) * 100,
    "%")
```

$\mathbf{P}(A_1 \mid \mathbf{B})$ **<-** $[\mathbf{P}(\mathbf{B} \mid A_1)$ * $\mathbf{P}(A_1)]$ / $[\mathbf{P}(\mathbf{B} \mid A_1)$ * $\mathbf{P}(A_1)] + [\mathbf{P}(\mathbf{B} \mid A_2)$ * $\mathbf{P}(A_1)]$

```
## [1] "The probability that an individual who is reported as positive by the new test actually has the
```

Key: Median cost per positive case (CPPC) = 100,000

Cost per Administration (CPA) = 1,000

Individuals (n) = 100,000

Prevalence Rate (pr) = .001

```r
# 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?

# Key
CPPC <- 1e+05
CPA <- 1000
n <- 1e+05
pr <- 0.001

# finding cost per case
cost_per_case <- CPPC + CPA
cost_per_case
```

```
## [1] 101000
```

```r
# Of the 100000 individuals, the prevalence rate of 0.1%,
# you will have 100 individuals that will get a positive
# result.
positive_case <- n * pr
positive_case
```

```
## [1] 100
```

```r
# With the 96% of actual positive cases from the 100
# individuals, only 96 will actually be positive and 4 will
# be false positive.

# Finding true negative and false negative cases.
true_negative <- (n - 100)
true_neg <- true_negative * 0.98  # multiplying by negative cases percent
true_neg
```

```
## [1] 97902
```

```r
false_negative <- true_negative - true_neg
false_negative
```

```
## [1] 1998
```

```r
# Find first year cost for positive cases
first_year_cost <- (96 + false_negative) * cost_per_case  # using the 96 of actual positive cases and a
paste("The total first year cost for treating 100,000 individuals is",
    first_year_cost)
```

```
## [1] "The total first year cost for treating 100,000 individuals is 211494000"
```

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?

Solution:

Key: n = size

P = probability

x = inspections

```r
# Setting variables
n <- 24
P <- 0.05
x <- 2

# What is the probability that, after 24 months, you
# received exactly 2 inspections?
paste("The probability that after 24 months you receive exactly 2 inspections is",
    round(dbinom(x, n, P), 4) * 100, "%.")
```

```
## [1] "The probability that after 24 months you receive exactly 2 inspections is 22.32 %."
```

```r
# What is the probability that, after 24 months, you
# received 2 or more inspections?  P(two or more) = 1 -
# P(1) - P(0)
paste("The probability that after 24 month you received 2 or more inspections is",
    round(1 - (dbinom(1, n, P) + dbinom(0, n, P)), 4) * 100,
    "%.")
```

```
## [1] "The probability that after 24 month you received 2 or more inspections is 33.92 %."
```

```r
# What is the probability that your received fewer than 2
# inspections?
paste("The probability that you received fewer than 2 inspections is",
    round(dbinom(1, n, P) + dbinom(0, n, P), 4) * 100, "%.")
```

```
## [1] "The probability that you received fewer than 2 inspections is 66.08 %."
```

```r
# What is the expected number of inspections you should
# have received?
paste("The expected number of inspections you should have received is",
    (n * P), ".")
```

```
## [1] "The expected number of inspections you should have received is 1.2 ."
```

```r
# What is the standard deviation?
paste("The standard deviation is", round(sqrt(n * P * (1 - P)),
    3), ".")
```

```
## [1] "The standard deviation is 1.068 ."
```

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?

Solution: Key:

r = rate = 10

```r
r <- 10

# What is the probability that exactly 3 arrive in one
# hour?
paste("The probability that exactly 3 arrive in one hour is",
    round(dpois(3, r), 4) * 100, "%.")
```

```
## [1] "The probability that exactly 3 arrive in one hour is 0.76 %."
```

```
# What is the probability that more than 10 arrive in one
# hour?
paste("The probability that exactly 3 arrive in one hour is",
    round(1 - ppois(10, r), 4) * 100, "%.")
```

```
## [1] "The probability that exactly 3 arrive in one hour is 41.7 %."
```

```
# How many would you expect to arrive in 8 hours?
per_hour <- 10
paste("The amount you would expect to arrive in 8 hours is",
    (8 * per_hour), "patients.")
```

```
## [1] "The amount you would expect to arrive in 8 hours is 80 patients."
```

```
# What is the standard deviation of the appropriate
# probability distribution?
paste("The standard deviation os the appropriate probability distribution is",
    round(sqrt(r)), ".")
```

```
## [1] "The standard deviation os the appropriate probability distribution is 3 ."
```

```
# 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?
paste("The percent utilization for three family practice providers that can see 24 templated patients ea
    round(80/72 * 100))
```

```
## [1] "The percent utilization for three family practice providers that can see 24 templated patients e
```

```
paste("The recommendations based on the amount expected versus the amount that can be served is to exte
```

```
## [1] "The recommendations based on the amount expected versus the amount that can be served is to exte
```

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
# Key
x <- 5
m <- 15
n <- 15
k <- 6

# What was the probability he/she would have selected five
# nurses for the trips?
paste("The probability he/she would have selected five nurses for the trips is",
    round(dhyper(x, m, n, k, log = FALSE), 4) * 100, "%.")
```

Solution:

```
## [1] "The probability he/she would have selected five nurses for the trips is 7.59 %."
```

```r
# How many nurses would we have expected your subordinate
# to send?
paste("The subordinate will need to send", 15 * 6/(15 + 15),
    "nurses.")
```

```
## [1] "The subordinate will need to send 3 nurses."
```

```r
# How many non-nurses would we have expected your
# subordinate to send?
paste("The subordinate will need to send", (6 - 3), "non-nurses.")
```

```
## [1] "The subordinate will need to send 3 non-nurses."
```

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
# Key
p <- 0.001
q <- 1 - p
h_12 <- 1200
h_15 <- 1500

# What is the probability that the driver will be seriously
# injured during the course of the year?
paste("The probability that the driver will be seriously injured during the course of the year is",
    round((pgeom(h_12, p)), 4) * 100, "%.")
```

Solution:

```
## [1] "The probability that the driver will be seriously injured during the course of the year is 69.93
```

```
# In the course of 15 months?
paste("The probabilty that the driver will be seriously injured during the course of 15 months is",
    round((pgeom(h_12 * (15/12), p)), 4) * 100, "%.")
```

```
## [1] "The probabilty that the driver will be seriously injured during the course of 15 months is 77.73
```

```
# What is the expected number of hours that a driver will
# drive before being seriously injured?
paste("The expected number of hours a driver will driver before being seriously injured is",
    (1/0.001), "hours.")
```

```
## [1] "The expected number of hours a driver will driver before being seriously injured is 1000 hours.
```

```
# Given that a driver has driven 1200 hours, what is the
# probability that he or she will be injured in the next
# 100 hours?
paste("The probability that he or she will be injured in the next 10 hours after driving 1200 hours is"
    round(pgeom(100, p), 4) * 100, "%.")
```

```
## [1] "The probability that he or she will be injured in the next 10 hours after driving 1200 hours is
```

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?

```
# What is the probability that the generator will fail more
# than twice in 1000 hours?
paste("The probability that the generator will fail more than twice in 1000 hours is",
    round(1 - ppois(2, 1), 4) * 100, "%.")
```

Solution:

```
## [1] "The probability that the generator will fail more than twice in 1000 hours is 8.03 %."
```

```
# What is the expected value?
paste("The expected value is", "lambda = 1")
```

```
## [1] "The expected value is lambda = 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?

```r
# Key
w <- 10
min <- 0
max <- 30

# What is the probability that this patient will wait more
# than 10 minutes?
paste("The probability that this patient will wait for more than 10 minutes is",
    round(1 - punif(w, min, max), 4) * 100, "%.")
```

Solution:

```
## [1] "The probability that this patient will wait for more than 10 minutes is 66.67 %."
```

```r
# 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?
PA <- 1 - punif(15, 10, 30)
PB <- 1 - punif(10, 10, 30)
paste("The probability that he / she will wait at least anohter 5 minutes prior to being seen is",
    (PA/PB) * 100, "%.")
```

```
## [1] "The probability that he / she will wait at least anohter 5 minutes prior to being seen is 75 %.
```

```r
# What is the expected waiting time?
paste("The expected waiting time is", 1/2 * (0 + 30), "minutes.")
```

```
## [1] "The expected waiting time is 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
lambda <- 1/10

# What is the expected failure time?
paste("The expected failure time is 10.")
```

Solution:

```
## [1] "The expected failure time is 10."
```

```r
# What is the standard deviation?
paste("The standard deviation is also 10.")
```

```
## [1] "The standard deviation is also 10."
```

```r
# What is the probability that your MRI will fail after 8
# years?
paste("The probability that your MRI will fail after 8 years is",
    round(1 - pexp(8, lambda), 4) * 100, "%.")
```

```
## [1] "The probability that your MRI will fail after 8 years is 44.93 %."
```

```r
# 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?
paste("The probability that it will fail in the next two years is",
    round(pexp((10 - 8), lambda), 4) * 100, "%.")
```

```
## [1] "The probability that it will fail in the next two years is 18.13 %."
```