Homework 5

Sam Reeves

9/24/2021

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.

```r
rate <- 0.001
n <- 100000

# Sensitivity
posGivInf <- 0.96

#Specificity
negGivNotInf <- 0.98

# Prevalence
totalInf <- n * rate
totalNotInf <- n - totalInf

negGivInf <- totalInf - (posGivInf * totalInf)
posGivNotInf <- totalNotInf * (1 - negGivNotInf)

population <- matrix(c(posGivInf*totalInf, negGivInf, totalInf,
                       posGivNotInf, negGivNotInf*totalNotInf, totalNotInf,
                       posGivInf*totalInf+posGivNotInf,
                       negGivInf+negGivNotInf*totalNotInf, NaN),
                     ncol = 3, byrow = TRUE)

colnames(population) <- c('positive', 'negative', 'total')
rownames(population) <- c('infected', 'not infected', 'total')
population
```

```
##              positive negative total
## infected           96        4   100
## not infected     1998    97902 99900
## total            2094    97906   NaN
```

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?

```
(infGivPos = posGivInf / (1-rate))
```

```
## [1] 0.960961
```

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?

```
cost.case <- 100000
cost.test <- 1000

infGivNeg <- negGivInf / (1-rate)
notInfGivPos <- posGivNotInf / (1-rate)

# plus test everybody
# plus people who are infected but negative are eventually tested again
# plus treatment for anybody who got a positive
# minus treatment for anybody who got a false positive
year <- cost.test * (n + infGivNeg) +
  cost.case * (posGivInf + posGivNotInf) -
  cost.case * (notInfGivPos)

year
```

```
## [1] 99900004
```

2. (Binomial)

The probability of your organization receiving a Joint Commission inspection in any given month is .05.

```
# Probability in months
p <- 0.05
```

What is the probability that, after 24 months, you received exactly 2 inspections?

```
n <- 24
x <- 2
dbinom(x, n, p)
```

```
## [1] 0.2232381
```

What is the probability that, after 24 months, you received 2 or more inspections?

```
q = 1
pLessThanTwo <- pbinom(q, size = n, prob = p)
(pTwoOrMore <- 1 - pLessThanTwo)
```

```
## [1] 0.3391827
```

What is the probability that your received fewer than 2 inspections?

```
pLessThanTwo
```

```
## [1] 0.6608173
```

What is the expected number of inspections you should have received?

```
# make a set of each case
x <- 0:24

# compute the p of each case
probs <- dbinom(x, n, p)

# subtract 1 because zero inspections is possible!
which.max(probs) - 1
```

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

What is the standard deviation?

```
sd(x * p)
```

```
## [1] 0.36799
```

3. (Poisson).

You are modeling the family practice clinic and notice that patients arrive at a rate of 10 per hour.

```
# patients per hour
lambda = 10
```

What is the probability that exactly 3 arrive in one hour?

```
# ppois is the Poisson function on a vec of quantiles
# lambda is a vec of non-negative means

# p(exactly3) = p(3) - p(2)
ppois(3, lambda) - ppois(2, lambda)
```

```
## [1] 0.007566655
```

What is the probability that more than 10 arrive in one hour?

```
q = 10
ppois(q, lambda, lower.tail = FALSE)
```

```
## [1] 0.4169602
```

How many would you expect to arrive in 8 hours? What is the standard deviation of the appropriate probability distribution?

```
t = 8

(muOfx <- t * lambda)
```

```
## [1] 80
```

```
(stanDevOfx <- sqrt(muOfx))
```

```
## [1] 8.944272
```

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?

```
patientsSeen <- 3 * 24
patientsArrived <- 10 * 8
(utilization <- patientsSeen / patientsArrived)
```

```
## [1] 0.9
```

I assume there is a daily cap on how many people will tend to arrive. Some sort of natural limit. So, I would experiment with opening for an extra hour in the morning or evening, and if demand is still not met, open a fourth location that is not situated in the geographical center of the clinics.

4. (Hypergeometric).

Your subordinate with 30 employees 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.

```
# nurses and non-nurses
m = 15
n = 15

# Total paid trips
k = 6

# Nurses on paid trips
x = 5
```

If your subordinate acted innocently, what was the probability he/she would have selected five nurses for the trips?

```r
(p5 <- dhyper(x, m, n, k, log = FALSE))
```

```
## [1] 0.07586207
```

How many nurses would we have expected your subordinate to send? How many non-nurses would we have expected your subordinate to send?

> There are equal amounts of nurses and non-nurses. The population is a reasonable size even number, and the number of people going on these trips is even. So, without calculating anything, we know it's got to be 3 nurses and 3 non-nurses.

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.

```r
# per hour
p = 0.001

# hours
n = 1200
```

What is the probability that the driver will be seriously injured during the course of the year?

```r
1 - (1 - p) ^ n
```

```
## [1] 0.6989866
```

In the course of 15 months?

```r
n = 1500
1 - (1 - p) ^ n
```

```
## [1] 0.7770372
```

What is the expected number of hours that a driver will drive before being seriously injured?

$$
\begin{aligned}
E(x) &= \sum_{x \in \Omega} x m(x) \\
&= \frac{1}{p}
\end{aligned}
$$

```r
1 / p
```

```
## [1] 1000
```

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
# We are only concerned with the future!
n = 100

1 - (1 - p) ^ n
```

```
## [1] 0.09520785
```

6. (Binomial).

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

```r
# per hour
p = 0.001
```

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

```r
n = 1000
q = 2
# P(more than two failures) = 1 - P(two or fewer failures)
1 - pbinom(q, n, p, lower.tail = TRUE, log.p = FALSE)
```

```
## [1] 0.08020934
```

What is the expected value?

```r
1000 / 1000
```

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

7. (Uniform).

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?

```r
possibilities <- length(0:30)
successes <- length(11:30)

successes / possibilities
```

```
## [1] 0.6451613
```

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
possibilities <- length(11:30)
successes <- length(15:30)

successes / possibilities
```

```
## [1] 0.8
```

Expected waiting time is the mean value of 15 minutes.

8. (Exponential).

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.

```r
rate <- 1 / 10
```

What is the expected failure time?

The manufacturer states that the expected lifetime is 10 years.

What is the standard deviation?

$$\mu_X = \frac{1}{\lambda}$$
$$\sigma_X^2 = \frac{1}{\lambda^2}$$
$$\mu_X = \sigma_X \ = \frac{1}{\lambda}$$
$$\sigma_X = 1$$

What is the probability that your MRI will fail after 8 years?

```r
q <- 8
e <- exp(1)

# P(8 or more) = 1 - P(less than 8)
# p <- 1 - (1 - e^(-rate * q))
(p <- e^(-rate * q))
```

```
## [1] 0.449329
```

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
#P(next 2) = P(10 years) - P(8 years)
p10 <- e^(-rate * 10)
p8 <- e^(-rate * 8)

1 - p10 - p8
```

```
## [1] 0.1827916
```