Homework 5

Kenan Sooklall

2/23/2021

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?

$P(+|HIV) = 0.96$
$P(-|HIV) = 0.98$
$P(HIV) = 0.001$

$P(HIV|+) = P(+|HIV) * P(HIV)/(P(+))$
$\rightarrow P(+|HIV) * P(HIV)/(P(+|HIV) * P(HIV) + P(+|!HIV) * P(!HIV))$

Plugging in the values

$0.96 * 0.001/(0.96 * 0.001 + 0.02 * 0.999) = 0.04584$

The probability that an individual who is reported as positive by the new test actually has ~50% of being positive with the disease. That high percentage is due to the fact that HIV is very rare 0.1%.

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?

With a population of 100000 and a prevalence rate of 0.1% that implies 100 individuals will get HIV. Of the 100 and a test sensitivity of 96% implies 96 of the 100 cases will be true positive that leaves 4 (100 - 96) false positive cases.

Similar 0.999 (1- 0.0001) of the population will be true negatives, that's 99900 individuals. A test specificity of 98% implies 97902 (99900 * 0.98) of the cases will be true negative leaving 1998 (99900 - 97902) false negatives.

```r
positives = c(96, 1998)
negatives = c(4, 97902)
median_cost = 101000
print(rbind(positives, negatives))
```

```
##            [,1]  [,2]
## positives    96  1998
## negatives     4 97902
```

```
total_cost = sum(positives) * median_cost
```

I assume it costs \$101,000 for each positive case in total, so the overall annual cost will be \$211,494,000

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?

```
p=0.05
n=24
k=2

dbinom(k, size=n, prob=p)
```

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

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

```
1 - (dbinom(1, size=n, prob=p) + dbinom(0, size=n, prob=p))
```

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

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

```
dbinom(1, size=n, prob=p) + dbinom(0, size=n, prob=p)
```

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

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

$\mu = 24 * p = 1.2$

What is the standard deviation?

$\sigma = sqrt(n * p * (1 - p)) = 1.07$

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? 0.7%

```
u = 10
x = 3

u^x * exp(-u) / (factorial(x))
```

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

What is the probability that more than 10 arrive in one hour? 41.6%

$P(X > 10) = 1 - P(X \leq 10)$

```r
u=10
x = c(0:10)

tmp = 0
for (i in x) {
  tmp = tmp + u^i/(factorial(i))
}
1 - exp(-u) * tmp
```

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

```r
1 - ppois(10, 10)
```

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

How many would you expect to arrive in 8 hours? 10/per hour * 8 hours = 80

What is the standard deviation of the appropriate probability distribution? 3.16

```r
sqrt(10)
```

```
## [1] 3.162278
```

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?

Three family practices that can see 24 patients each day implies 72 (24 * 3) templated patients per day. Assuming the practice is open on an 8 hr day there would be (8 * u) 80 patients a day meaning the practice will be understaffed (80/72=1.11); however adding another family would be overkill (24 * 4=96) person for a 8 hour day. However, I would recommend 4 families that can see 96 patients at a practice that is open 9 hours a day. So 90 patients a day with staff that can handle 96 patients (90/96) 93.75% utilization.

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?

```r
choose(15,5) * choose(15,1) / (choose(30, 6)) +choose(15,6) * choose(15,0) / (choose(30, 6))
```

```
## [1] 0.08429119
```

```r
1 - phyper(4,15,15,6)
```

```
## [1] 0.08429119
```

There is a 8% chance my subordinate acted innocently

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

```
15 * 6 / (15 + 15)
```

```
## [1] 3
```

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

```
15 * 6 / (15 + 15)
```

```
## [1] 3
```

Problem 5 - Geometric

$P(X = x) = q^{x-1}p$
mean = 1/p

var = (1/p)*(1/p-1)

where $p - success$
$q - fail \rightarrow q = 1 - p$

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?

```
p = 0.001
q = 1-p
x=1200
not_being_injured = q^x
injured = 1 - not_being_injured
injured
```

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

In the course of 15 months?

```
1 - q^1500
```

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

```
pgeom(1500, p)
```

```
## [1] 0.7772602
```

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

```
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
(1-p)^99 * p * choose(100,1)
```

```
## [1] 0.09056978
```

```r
pgeom(100, p)
```

```
## [1] 0.09611265
```

I am not sure why those two approaches are different. The probability that he or she will be injured in the next 100 hours = ~9.6%

Problem 6 - Poisson

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?

```r
1 - ppois(2,1)
```

```
## [1] 0.0803014
```

What is the expected value?

$\lambda = 1$

Problem 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?

$P(X > 10) = 1 - P(X <= 9) \rightarrow 1 - 10/30 = 0.67$

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?

$1 - (15 - 10)/(30 - 10) \rightarrow 75.0$

What is the expected waiting time?

$\mu = (a + b)/2 \rightarrow 30/2 = 15$
$\sigma = sqrt((b - a)^2/12) \rightarrow sqrt(30^2/12) = 8.66$

Problem 8 - Exponential

$F(x) = \lambda * e^{(-\lambda * X)}$

$\mu = \sigma = 1/\lambda$

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? 10 years

What is the standard deviation? 10

What is the probability that your MRI will fail after 8 years? P(X>8) = 1 - P(X<=8)

```
lambda = 0.1
1 - pexp(8, lambda)
```

```
## [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?

```
(pexp(10,lambda) - pexp(8,lambda))/(1 - pexp(8,lambda))
```

```
## [1] 0.1812692
```

```
pexp(2,lambda)
```

```
## [1] 0.1812692
```