Week 5

Shaya Engelman

2024-02-22

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?

Using a contingency table easily showed that the

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

Now, for the total first-year cost for treating 100,000 individuals:

Total cost = (Number of individuals) * (Cost per positive case + Cost of the test) Number of individuals = 100,000 Cost per positive case = 100,000 Cost of the test = 1000

Plug those numbers in and you get **Total Cost = 10,100,000,000**

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?

I used R for these answers, the answers are printed below (as far as expected number of inspections, since yo cannot have a partial inspection we should round that down to 1):

```
p_exactly_2 <- dbinom(2, size = 24, prob = 0.05)

p_2_or_more <- 1 - pbinom(1, size = 24, prob = 0.05)

p_less_than_2 <- pbinom(1, size = 24, prob = 0.05)

expected_value <- 24 * 0.05

standard_deviation <- sqrt(24 * 0.05 * 0.95)
```

```r
cat("Probability of exactly 2 inspections:", p_exactly_2, "\n")
```

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

```r
cat("Probability of 2 or more inspections:", p_2_or_more, "\n")
```

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

```r
cat("Probability of fewer than 2 inspections:", p_less_than_2, "\n")
```

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

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

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

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

```
## 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?

```r
lambda <- 10  # average rate of arrival per hour
time_period <- 8
providers <- 3
patients_per_provider <- 24
total_capacity <- providers * patients_per_provider

prob_3_arrivals <- dpois(3, lambda)

prob_more_than_10_arrivals <- 1 - ppois(10, lambda)
#The following was just to test it, I wan't sure how it goes upwards of 10 to infinity
#prob_more_than_10_arrivals_test <- 1 - sum(dpois(0:10, lambda))
#all.equal(prob_more_than_10_arrivals_test, prob_more_than_10_arrivals, tolerance = 1e-10)


expected_arrivals <- lambda * time_period

standard_deviation <- sqrt(lambda * time_period)

percent_utilization <- (lambda / total_capacity) * 100

cat("Probability of exactly 3 arrivals in one hour:", prob_3_arrivals, "\n")
```

```
## Probability of exactly 3 arrivals in one hour: 0.007566655
```

```
cat("Probability of more than 10 arrivals in one hour:", prob_more_than_10_arrivals, "\n")
```

```
## Probability of more than 10 arrivals in one hour: 0.4169602
```

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

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

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

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

```
cat("Percent utilization:", percent_utilization)
```

```
## Percent utilization: 13.88889
```

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?

Assuming innocence and random selection, the split would be expected to be 3 nurses and 3 non-nurses. As far as the probability of selecting 5 nurses, see the following code.

```
workers <- 30
nurses <- 15
trip <- 6
trip_nurses <- 5

prob <- dhyper(trip_nurses, nurses, workers - nurses, trip)

expected_value <- lambda

cat("Expected value:", expected_value, "\n")
```

```
## Expected value: 10
```

```
cat("Probability of selecting exactly 5 nurses:", prob)
```

```
## Probability of selecting exactly 5 nurses: 0.07586207
```

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?

The answers are below. The probability of being injured within the next 100 hours are unchanged by the fact that the driver has already driven 1200 hours uninjured (this is assuming we don't recalculate the danger rate using this new information).

```
prob_injured_one_hour <- 0.001

prob_not_injured_one_hour <- 1 - prob_injured_one_hour

one_month <- 1200 / 12

prob_not_injured_12_months <- prob_not_injured_one_hour ^ (12 * one_month)

prob_injured_12_months <- 1 - prob_not_injured_12_months

prob_injured_15_months <- 1 - (prob_not_injured_one_hour ^ (one_month * 15))

expected_hours_before_injured <- 1 / prob_injured_one_hour

prob_injured_next_100_hours <- 1 - (prob_not_injured_one_hour ^ 100 ) #equivalent to one month

cat("Probability of being injured in one year:", prob_injured_12_months, "\n")
```

```
## Probability of being injured in one year: 0.6989866
```

```
cat("Probability of being injured in 15 months:", prob_injured_15_months, "\n")
```

```
## Probability of being injured in 15 months: 0.7770372
```

```
cat("Expected number of hours before being injured:", expected_hours_before_injured, "\n")
```

```
## Expected number of hours before being injured: 1000
```

```
cat("Probability of being injured in the next 100 hours (one month):", prob_injured_next_100_hours)
```

```
## Probability of being injured in the next 100 hours (one month): 0.09520785
```

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?

Answers below. I printed the expected value of the probability. For the expected amount over 1000, or any given amount of, hours, multiply the lambda by the hours)

```r
lambda <- 10  # average rate of arrival per hour

prob_less_than_or_equal_2_failures <- ppois(2, lambda)

prob_more_than_2_failures <- 1 - prob_less_than_or_equal_2_failures

cat("Probability of generator failing more than twice:", prob_more_than_2_failures, "\n")
```

```
## Probability of generator failing more than twice: 0.9972306
```

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
lower <- 0
upper <- 30

prob_more_than_10_minutes <- (upper - 10) / (upper - lower)

# Probability of waiting at least another 5 minutes
prob_at_least_another_5_minutes <- (upper - (10 + 5)) / (upper - 10)

# Calculate the expected waiting time
expected_waiting_time <- (lower + upper) / 2

# Print the result
cat("Probability of waiting more than 10 minutes:", prob_more_than_10_minutes, "\n")
```

```
## Probability of waiting more than 10 minutes: 0.6666667
```

```r
cat("Probability of waiting at least another 5 minutes after already waiting for 10 minutes:", prob_at_
```

```
## Probability of waiting at least another 5 minutes after already waiting for 10 minutes: 0.75
```

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

```
## 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. 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?

In an exponential distribution. the standard deviation is the same as mean rate. I calculated the probability that the machine will have failed after 8 years. If the question meant to ask for the probability that the machine will only fail after 8 hours (meaning, not before that), remove the "1 -" from the calculation to get its inverse. Also, an exponential distribution is considered memoryless. Meaning, the likelihood of the event happening within a specific timeframe doesn't get greater with time passing. However, life expectancy, (or machine life), is typically not accurately measured like that. My result below prints the probability assuming a true exponential distribution, for this specific case, a different model is probably more accurate.

```
expected_lifetime <- 10

lambda <- 1 / expected_lifetime

expected_failure_time <- 1 / lambda

standard_deviation <- 1 / lambda

prob_failure_by_8_years <- pexp(8, lambda)

prob_failure_next_2_years <- pexp(2, lambda)

cat("1. Expected failure time:", expected_failure_time, "years\n")
```

```
## 1. Expected failure time: 10 years
```

```
cat("2. Standard deviation:", standard_deviation, "years\n")
```

```
## 2. Standard deviation: 10 years
```

```
cat("3. Probability of failure by 8 years:", prob_failure_by_8_years, "\n")
```

```
## 3. Probability of failure by 8 years: 0.550671
```

```
cat("4. Probability of failure in the next two years given ownership for 8 years:", prob_failure_next_2
```

```
## 4. Probability of failure in the next two years given ownership for 8 years: 0.1812692
```