Review of MATH/STAT 394

Practice problems

Autumn 2024

1. In a very popular Massive Open Online Course (MOOC), there are two tests that students can use to check their progress. The tests aim to examine quite different skills and knowledge, and all students continue in the MOOC regardless of their test outcomes. In test-1, 60% of the students achieve a passing score. In test-2, 70% of the students achieve a passing score.

a. If the outcomes on the two tests are independent, what percentage of students pass both tests?

b. Suppose in fact, 50% of students pass both tests. What is the probability that a student

   i. will fail both tests?

   ii. who passes the first test will also pass the second test?

   iii. who passes the second test also passed the first test?

2. In a certain large population, there is a new flu virus. Each individual is susceptible to the virus with probability 1/3, and not susceptible with probability 2/3. In any given winter, a susceptible individual gets this type of flu with probability 3/4, and a non-susceptible individual gets this flu with probability 1/4.

a. Suppose Fred represents a randomly selected individual from this population. What is the probability that he gets the flu this winter?

b. Fred gets the flu. What is the probability that Fred was a susceptible individual?

3. Justin wants to know if his cat, Gus, prefers his right paw or if he uses both paws equally. So he dangles a ribbon in front of Gus and notes which paw Gus uses to bat at it.

   He does this 10 times and Gus bats at it with his right paw 8 times and his left paw 2 times. Then Gus gets bored and leaves.

   Let the random variable X denote the number of times that Gus batted with his right paw in 10 trials. (We observe $x = 8$). Suppose you decide to model X as a binomial random variable with $\pi = 0.5$ (meaning Gus is equally likely to use either paw).

a. What assumptions do you need to make in order for the binomial model to be a reasonable choice for this setting?

b. Give an expression for $P(X \geq 8)$. How will you calculate this using R? Go ahead and log on to the Rstudio image in JupyterHub (see link in navigation pane of course CANVAS) and calculate it.

4. A condition C among new-born babies occurs apparently independently in any baby with probability $p = 0.15$. Suppose in King County in 2014, there are 10,000 births.

a. Which probability distribution provides the most accurate model for computing the probability that more than 1600 babies with condition C are born in the county K in 2014? Explain your thinking. Go ahead and calculate it in R.

b. What other probability distribution might be more computationally convenient, and would provide a good approximation for the probability in part a? Justify your thinking. Go ahead and calculate it in R.

Hint: did you learn the Poisson approximation for the binomial?

5. Karen is studying for a history exam, where the teacher is going to choose 5 essay questions randomly from the 10 he has given the class. Due to an upcoming probability exam, she only has limited time to prepare for the history exam. Suppose she decides to study 3 out of the 10 questions.

   Calculate the probability that Karen will not have studied any of the questions on the exam.

6. In a certain country commercial airplane crashes occur according to a Poisson process at the rate of 2.5 per year. Using R, find the probability that the next two crashes will occur more than three months apart.

   Hint: Let X denote the number of crashes that occur in the next three months. Then $X \sim Pois\left(\lambda = 2.5 \times \frac{1}{4}\right)$

7. The annual rainfall (in inches) in a certain region is normally distributed with $\mu = 40$ and $\sigma = 4$. What is the probability that, starting with this year, it will take over 10 years before a year occurs having a rainfall of over 50 inches? What assumption are you making? You are given that `pnorm(2.5) = 0.994`

   Hint: Let X denote the number of years before a year occurs having rainfall of over 50 inches. What is a reasonable distribution to assume for X?

8. Suppose $X \sim Exp(1)$, that is it has PDF

$$f_X(x) = e^{-x} \qquad 0 \leq x < \infty$$

   What distribution does the random variable $Y = \frac{X}{\lambda}$ have? State the name of the distribution and also the value for any parameters. (Hint: Find the CDF of Y and then differentiate it to find a PDF)

9. Which distribution has the smaller 25th percentile? The $Unif(0, 1)$ or the $Exp(1)$?

10. The internal temperature in a gizmo is a random variable X with PDF (in appropriate units)

$$f(x) = \begin{cases} 11(1-x)^{10} & 0 < x < 1 \\ 0 & otherwise \end{cases}$$

   The gizmo has a cutoff feature, so that whenever the temperature exceeds the cutoff (call it k), the gizmo turns off. It is observed that the gizmo shuts off with probability 10^{-22}. What is k?