Stat 251: Assignment 4

Fall 2021

1. (This problem focuses on comparing proportions from two independent populations.) In a 2014 article titled "Gender Differences in Willingness to Guess," the author describes performing an experiment to determine if men or women are more likely to skip questions on the History SAT II test. (The author also discusses the implications on grade when more questions are skipped.) We will use a portion of the data from this experiment for this homework question.

   In an SAT-like environment, each student was given 20 questions and told that they would earn 1 point for every question answered correctly and lose 0.25 points for every question answered incorrectly. Questions left blank would earn (and lose) 0 points. We will treat each individual and question as independent. The table below shows the total number of questions asked with the total number skipped for each gender. (There were 85 women and 63 men in the study.)

Gender	Total Number of Questions	Number of Skipped Questions
Men	1260	67
Women	1700	173

   a. Let θ_W be the proportion of questions women test-takers skip and θ_M be the proportion of questions men test-takers skip. Assume *a priori* that both genders are more likely to answer the question than not by letting both parameters follow a Beta(0.5, 3) distribution. Plot the prior distribution. Report the prior probability that the proportion of skipped questions (for either gender) is less than 10%.

   b. Report the posterior distributions for θ_W and θ_M.

   c. What is the posterior probability that women will skip less than 10% of questions? And for men? Comment on how these compare to the prior probability.

   d. Plot the posterior distribution on the difference between the proportion of questions skipped by women and the proportion skipped by men. Use this distribution to answer the question of whether women will skip more questions than men.

2. (This problem focuses on learning about the Poisson distribution.) Let $X|\lambda \sim \text{Pois}(\lambda)$.
   a. If $\text{SD}(X|\lambda) = 2.2$, what is the value of λ?
   b. What is $P(X \leq 3|\lambda = 1)$? $P(X \leq 3|\lambda = 10)$? $P(X \leq 3|\lambda = 100)$?
   c. Plot the probability distributions $\lambda = 1$, $\lambda = 10$, and $\lambda = 100$. Comment on how the probability distributions are similar and different.
   d. When $\lambda = 4$, what is the 95th percentile?
   e. If $x_{.95}$ represents the 95th percentile you computed in d, find $P(X \leq x_{.95}|\lambda = 4)$. Why is it not equal to 0.95?

3. (This problem focuses on learning about the gamma distribution.) Let $\lambda \sim \text{Gamma}(\gamma, \phi)$.
   a. γ is known as the shape parameter. Plot the gamma distribution for two different values of γ and fixing $\phi = 3$. How does the shape of the distribution change for the different values of γ?
   b. ϕ is known as the rate parameter. Plot the gamma distribution for two different values of ϕ and fixing $\gamma = 5$. How does the distribution change in this case?
   c. Find $P(2 \leq \lambda \leq 5|\gamma = 2, \phi = 1.5)$.
   d. Find a 95% probability interval for $\lambda \sim \text{Gamma}(\gamma = 2, \phi = 1.5)$.