

Section 4

With questions by Chris

1 Beta \pm

You have observed 30 successes and 20 fails for an event with unknown probability. Based on this information you can model the probability of success, X , as a Beta random variable. You want to make a claim of the form: “the probability of success is $m \pm b$.”

Select m to be the expectation of X . Select b to be the smallest value, *rounded to two decimal places*, such that $P(m - b < X < m + b)$ is greater than or equal to 95%. Provide your answer as pseudocode that prints out the values m and b .

2 Bayesian RNA Quantification

Let N_A and N_B be the count of type-A RNA and type-B RNA in a person’s body. Both are Poisson distributed:

$$N_A \sim \text{Poi}(12 \cdot X) \quad N_B \sim \text{Poi}(12 \cdot (1 - X))$$

N_A and N_B have rates that come from an unknown continuous variable X . X takes on values between 0 and 1 and is an indicator of a person’s overall health. N_A and N_B are independent, conditioned on knowing X .

- a. For part (a) only, assume $X = 1/4$. What is $P(N_A > 10)$?
- b. Our prior belief is that $X \sim \text{Beta}(a = 2, b = 2)$. An experiment observes $N_A = 12$ and $N_B = 16$. Based on these observations, what is your updated distribution for X ? In other words, what is $f(X = x | N_A = 12, N_B = 16)$? You can leave your answer in terms of a normalization constant.
- c. What is the variance of your updated belief distribution for X , found in part (b)?

3 Midterm Review (Optional)

m strings are hashed (not necessarily uniformly) into a hash table with n buckets. Each string's hash is an independent trial, and the probability that a string hashes to bucket i is p_i , with $\sum_{i=1}^n p_i = 1$.

- a. Let E be the event that bucket 1 has ≥ 1 string hashed into it. What is $P(E)$?
- b. Let E be the event that at least one of buckets 1 to k has ≥ 1 string hashed into it. What is $P(E)$?
- c. $E =$ each of buckets 1 to k has ≥ 1 string hashed into it. What is $P(E)$?