SuperAGI Assignment

By

Bhavneek Singh

2020EE10482

IIT Delhi

Q/A

1. W_n = W_new(n+1)
2. Looking up a Z score table, the Z score has to be either < -1.96 or > 1.96 for us to be 95% confident about the conclusion.

Z score for 2 population proportions =

(p1 - p2)/

(p (1-p) (1/n1 + 1/n2))*(0.5)

Here, p1 = CTR in template 1

p2 = CTR in template 2

p = pooled proportion = (p1 * n1 + p2 * n2)/(n1 + n2)

n1 = number of people in population 1

n2 = number of people in population 2

Filling in the numbers, I find that the Z scores for

(A,B) = 2.4054.

(A,C) = 1.1577

(A, D) = -1.4293.

(A, E) = -2.7524

Comparing these Z scores for the bounds of a 95% confidence, we can see that E and B are confident conclusions, whereas we need more data for being more confident in decisions for C and D. Thus, option 2 is correct, 2. E is better than A with over 95% confidence, B is worse than A with over 95% confidence. You need to run the test for longer to tell where C and D compare to A with 95% confidence.

4. I believe that method 3 is the best approach in terms of improving accuracy. It seeks to train the model V2 on samples that V1 is performing poorly on, thus it is most likely to improve the weaknesses of V1 that could not be learnt from the 10000 samples of NY Times.

On the other hand, Method 2 is agnostic to V1's weakness, and method 1 is only works on samples that were anyways very nearly correctly predicted. Ranking for accuracy, I believe the models will rank M3 > M2 > M1 in decreasing accuracy. Note that this analysis rests on the fact that the cost and effort differences in obtaining these samples and labels is ignored.

5. a) MLE Estimation : Unfair Coin toss is well modelled by Bernoulli trials: Likelihood of obtaining k heads in n tosses is nCk (p)^k (1-p)^(n-k). There is also a digression into the theoretical results of Beta Distributions here. derivative(Log-likelihood) = 0 => k/p = (n-k)/(1-p). => p = k/n estimate.

b) prior distbn p(theta) = 1 for theta in [0,1] = beta(1,1)

posterior probabilities = beta(k+1, n-k+ 1)

Expected value of a beta distribution = alpha / (alpha + beta)

Thus, p = (k+1) / (n + 2)

c) The uniform distribution prior is equivalent to a beta(1,1) distbn. After observing k heads and n-k tails, we know that the beta distribution of belief changes to beta(k + 1, n-k + 1) . The mode of a beta distribution happens at (alpha-1)/(alpha + beta - 2). This, the MAP estimate is (k+1 - 1) / (k+1 + n-k + 1 -2) = k/n, which is the same as MLE!