

WSC15 paper

June 21, 2015

Reviewer 1.

1. *Frazier 2014 was cite lots of times, modify the cite commands accordingly to make sure the correct format.* Now, it should be fine.
2. *Second corrections.* They're now correct.
3. *Use Section 2, 3, 4 instead just 2,3,4 when referring to those sections.* I changed this part, but I'm not sure if we have to write the word section before the number.
4. Page 3, 3rd paragraph. Now it is correct.
5. Page 3, 4th paragraph. Now it is correct.
6. Page 3, 1st paragraph in Section 3. Now it is correct.
7. 3rd paragraph in Section 3. I think that now it is clear. *Says 100 is the recommended value for n_0 , but later in the algorithm and experiments used something else.* I added that this value is recommended when the variances are unknown.
8. End of page 3. Now it should be clear.
9. Section 4, 3rd paragraph. Now it is correct.
10. Section 6. Now it is correct.

Reviewer 2.

1. *It seems that the choice of n_0 is relatively arbitrary. In practice, can problem properties be used to obtain some understanding of how large/small n_0 should be?* I don't now if we should say something in section 3 about this because it is difficult to say something if we don't now the variances of the alternatives.
2. *It does seem to me that we choose a far too stringent threshold, the resulting problem may be rendered infeasible since μ/k are exogenously specified and presumably unavailable a prior and δ is a user-specified parameter. In re-reading the paper, it seems that the alternatives may have nearly the same means, in which case, the likelihood of infeasibility may be modest.* Should we say that if the configuration is not in the preference zone, then we can't say anything about the alternative selected?

3. *While the BIZ procedure is asymptotically valid, naturally in practice, the schemes have to be terminated after a finite number of samples. Are there error bounds available, even in a restricted regime, that can govern the choice of when the procedure may be terminated.* I said in section 3 that the algorithm will end in a finite number of steps. However, I think there are no error bounds available if we just stop the algorithm before time. What should we say about this?
4. *In the numerical experiments, the choice of μ 's and δ ensures feasibility. Suppose μ_{100} is set to $\delta + 1$, then unless I have misunderstood how the scheme works, the threshold constraint would render the problem infeasible. In this case, how would the scheme respond?* This can't happen because we are supposing that $\mu = a\delta$. Should we say something else?
5. *It might be useful to see how the BIZ procedure is indeed less conservative in practice compared to its counterparts in the R&S literature. The authors claim that the proposed procedure requires less samples – presumably this is only an empirical observation. If so, please emphasize that this is the case; else it might be worth pointing to the formal statement that guarantees this claim.* This observation is based on empirical observations and that the BIZ procedure is less conservative. Should we say something else in the introduction to make this more clear?
6. *Finally, some quick observations regarding the references.* I think that now they are fine. I had to cite the two editions of Billingsley's book because I used the two editions.