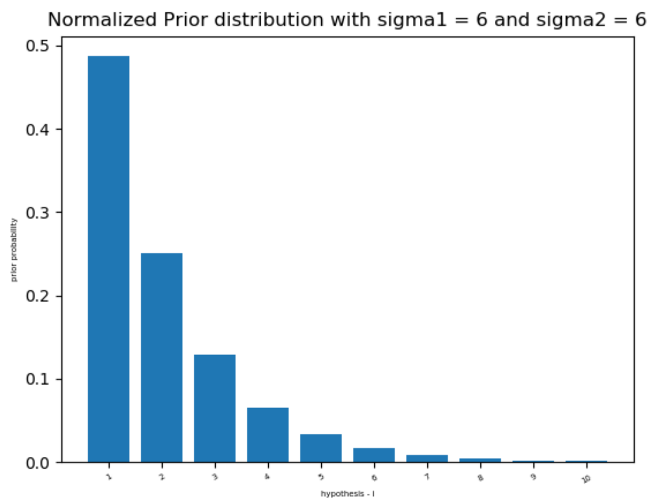
CSCI-5822 Probabilistic Models of Human and Machine Intelligence
Kumar Bhargav Srinivasan **kusr7198@colorado.edu**

Assignment 1

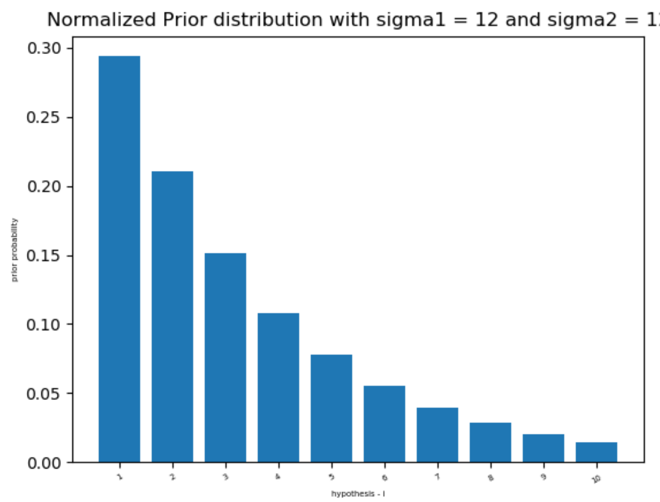
1st February 2018

---

1. Task 1 Prior
   Make a bar graph of the prior distribution, $P(H)$, for $\sigma 1 = \sigma 2 = 6$. Make a graph of the prior distribution for $\sigma 1 = \sigma 2 = 12$.
   Using formula for expected size prior : $P(H) = \exp(-(s1/\ \sigma 1 + s2/\ \sigma 2))$
   following is the bar graph for $\sigma = 6$ and $\sigma = 12$



(a) $\sigma = 6$



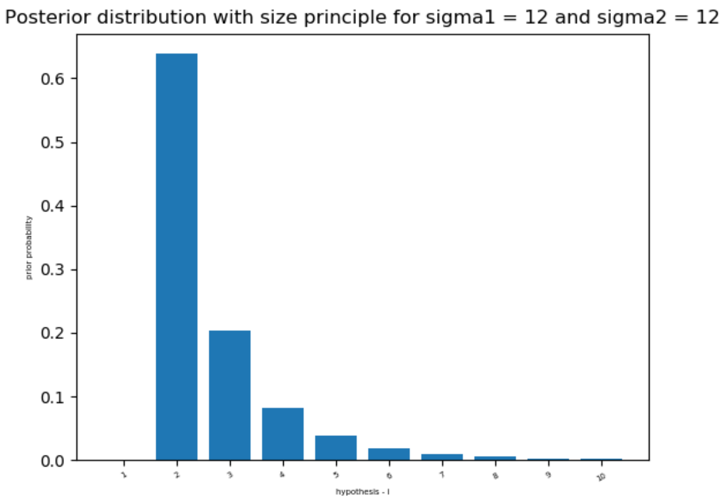
(b) $\sigma = 12$

Figure 1: task1 Prior probability for sigma = 6 and 12

Hypothesis size	sigma = 6	sigma = 12
1	0.48720291186115366	0.29395524695030584
2	0.2501383153920434	0.21062813834734448
3	0.12842529324824878	0.15092165601374014
4	0.06593574407043169	0.10814009197749183
5	0.03385253976191134	0.07748576183020232
6	0.01738047343649679	0.05552097447500338
7	0.008923432599188675	0.039782516604908065
8	0.004581443056956966	0.028505418760838314
9	0.0023521912953147167	0.020425025063158155
10	0.0012076552782540222	0.014635169977007203

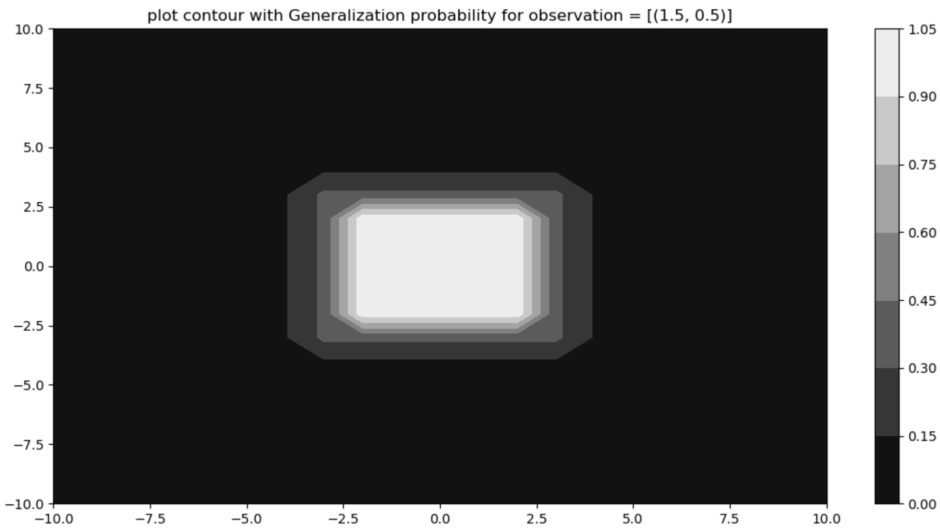
Following table shows the difference in priors for given hypotheis space for different σ
Large the sigma value, more Uniform the prior (more distributed the prior)

2. Task 2 Posterior
   With observation, X = (1.5, 0.5), compute the posterior $P(H|X)$ with $\sigma1 = \sigma2 = 12$.
   Following is the posterior probabilities for hypothesis sizes:
   1: 0.0
   2: 0.6384037371483289
   3: 0.203305007312932
   4: 0.0819418518765838
   5: 0.03757689761022202
   6: 0.018697933119409354
   7: 0.009843174751644325
   8: 0.005399909411642571
   9: 0.0030571489720130295
   10: 0.001774339797224005

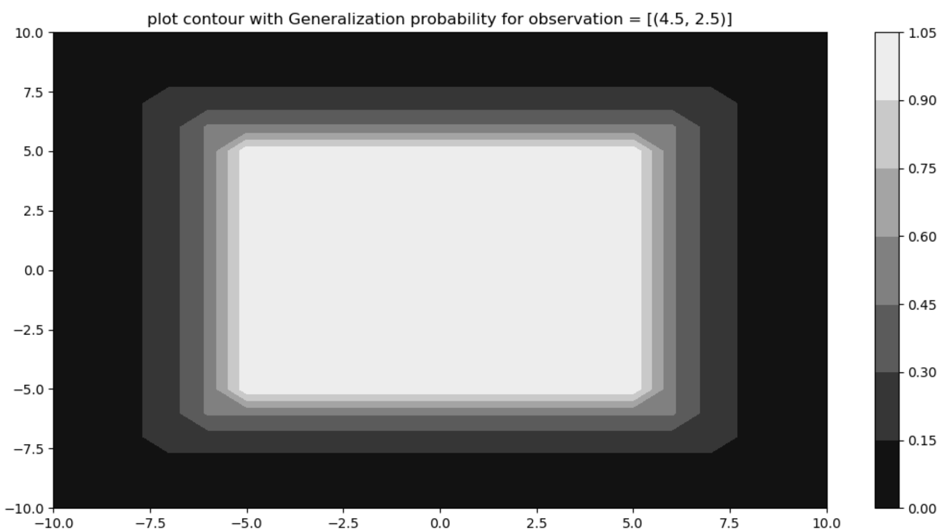


   Posterior are calculated by finding likelihood using size principle: $P(X|H) = 1/|h|^n$ if $\forall j, x^j$ in h, 0 otherwise and multiplying with priors

3. Task 3 Generalization
   compute generalization predictions, $P(y \rightarrow concept|X)$, over the whole space of possible generalization points, y, for $X = (1.5, 0.5)$ and $\sigma1 = \sigma2 = 10$.
   Below Figure discribes a contour of generalizations from -10 to 10 accross all the points of distance 1 unit between each. Generalization probabilities is higher near the hypothesis space which has higher posterior.

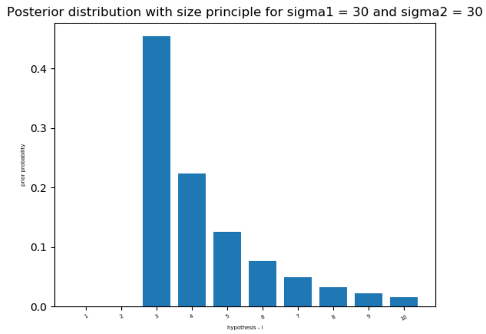


4. Task 4 : With observation X = (4.5, 2.5), compute the posterior $P(H|X)$ with $\sigma 1 = \sigma 2 = 12$.

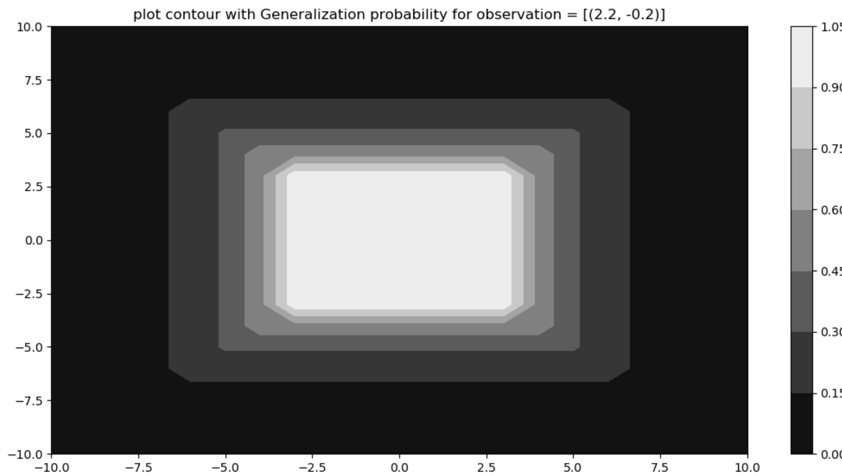


Same as in task 4 for a different observation

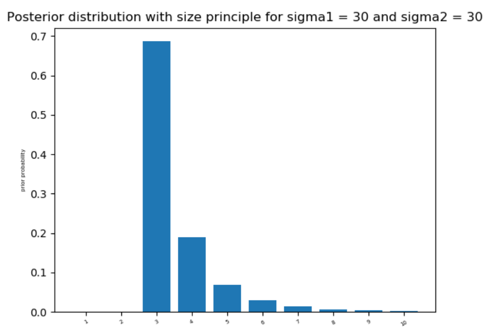
5. Task 5 Generalization
Infering the plots below, As the points for the given hypothesis increases,the denominator : $math.pow(4 * s1 * s2, len(self.observation))$ increases for all the hypothesis where the values belong and the product of the prior and likelihood is significant only for these hypothesis.Though absolute posterior values are lower, On normalizing the values over all the instances, posterior becomes significant for valid hypothesis.
for the hypothesis where all the points won't belong likelihood become 0 and for those where all belong will have lower likelihood.Product of the prior and likelihood is lower, but on normalizing these values, posterior increases. So as the points increase the Likelihood overcomes the dominance of prior.
Also, generalization probabilities become significant around the points observed due to the above fact of posterior being high.



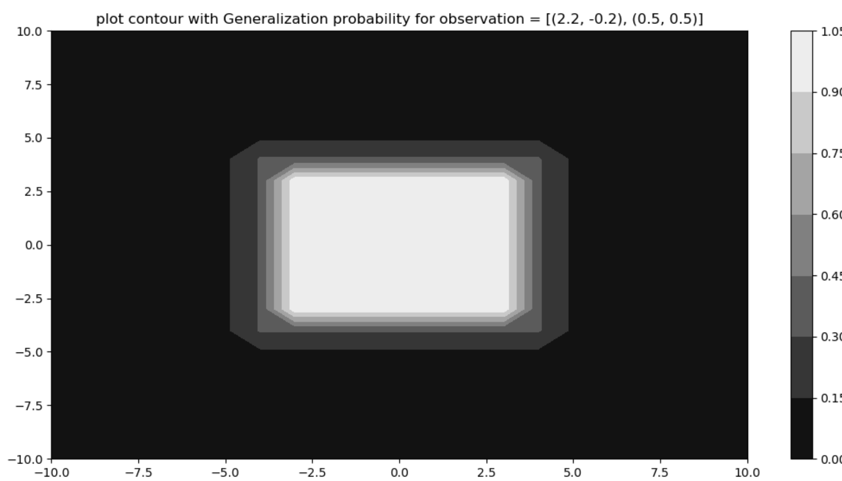
(a) $X = (2.2, -.2)$



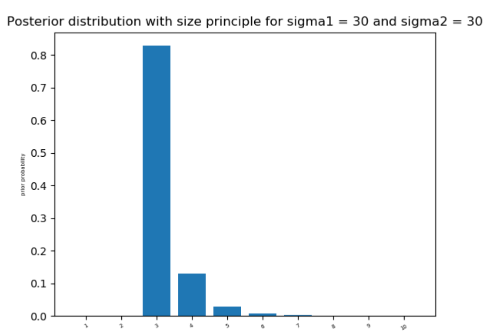
(b) $X = (2.2, -.2)$



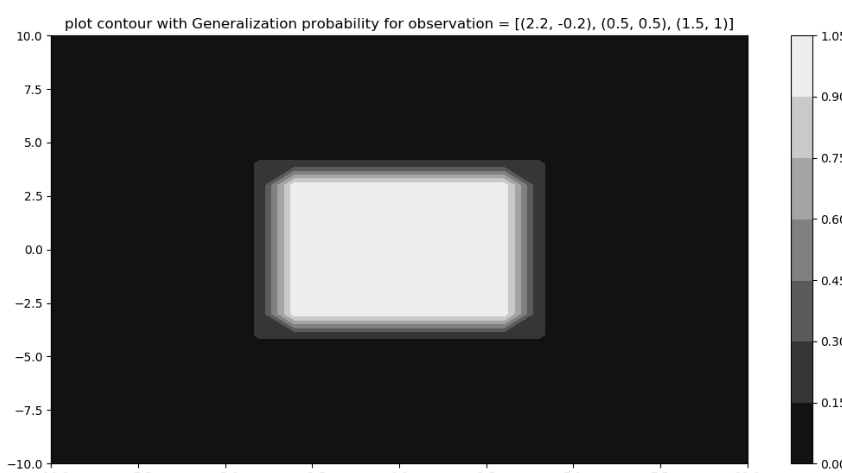
(c) $X = (2.2, -.2), (.5, .5)$



(d) $X = (2.2, -.2), (.5, .5)$



(e) $X = (2.2, -.2), (.5, .5), (1.5, 1)$

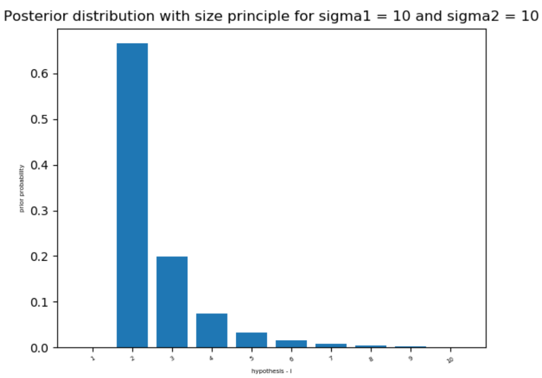


(f) $X = (2.2, -.2), (.5, .5), (1.5, 1)$

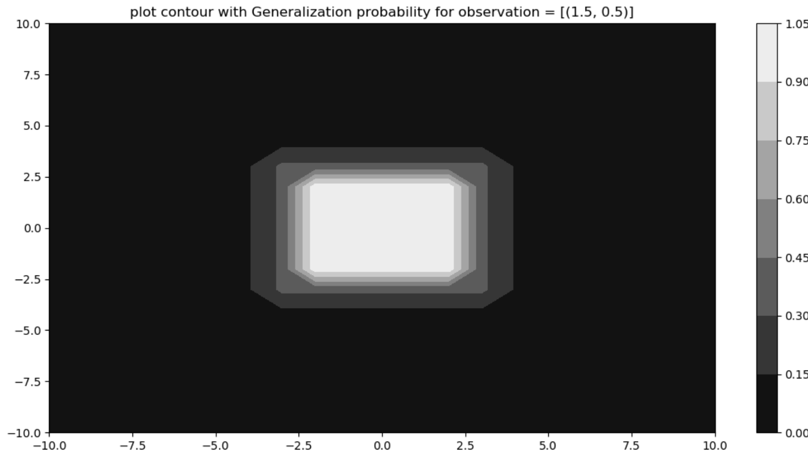
Figure 2: Posterior probability and generalization for $X = (2.2, -.2), X = (2.2, -.2), (.5, .5)$, and $X = (2.2, -.2), (.5, .5), (1.5, 1)$

6. Task 6 Generalization

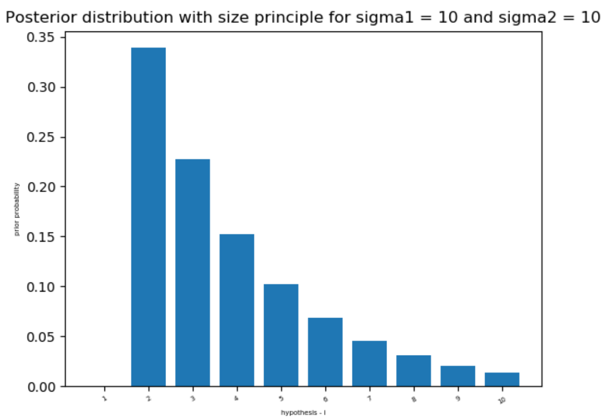
   (a) **With and without size principle**, Without size principle we treat likelihood as 1 for all the hypothesis where points belong to, So posterior is still low when compared to size principle, Hence couldn't converge to the target hypothesis with sampler sample size.



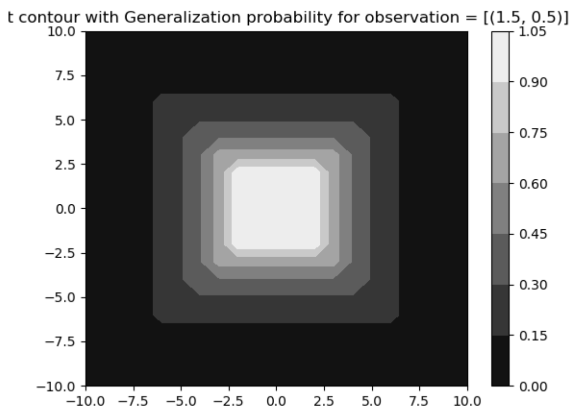
(a) expected size-posterior



(b) expected size-generalization



(c) without size principle–posterior

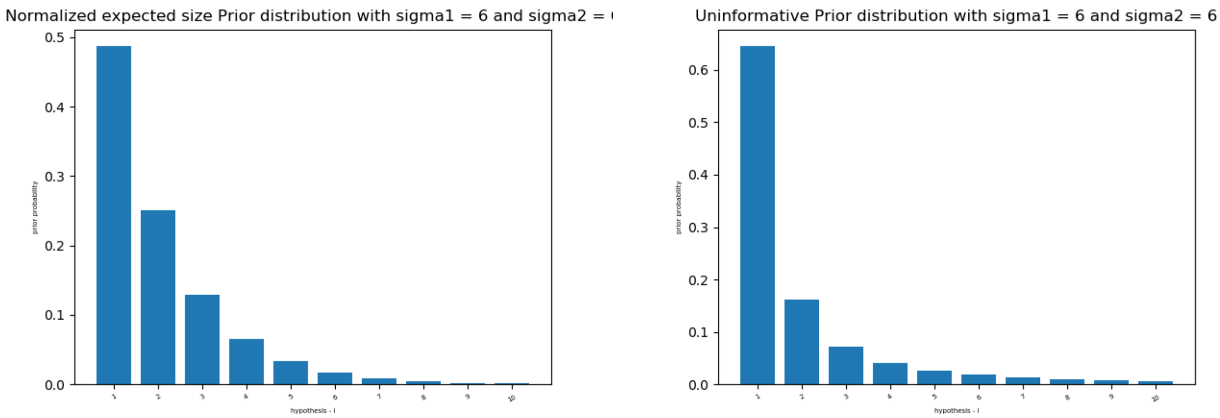


(d) without size principle-generalization

Figure 3: Posterior probability and generalization for $X = (1.5, 0.5)$

(b) **uniform vs expected-size prior**

Uniform prior tends to increase posterior drastically on increase of the number of samples, when compared to the expected-size prior, so if there is any inherent knowledge of the hypothesis, uniform prior tends to fail whereas expected-size prior tends to accomodate it.
Comparision of priors on sigma = 6



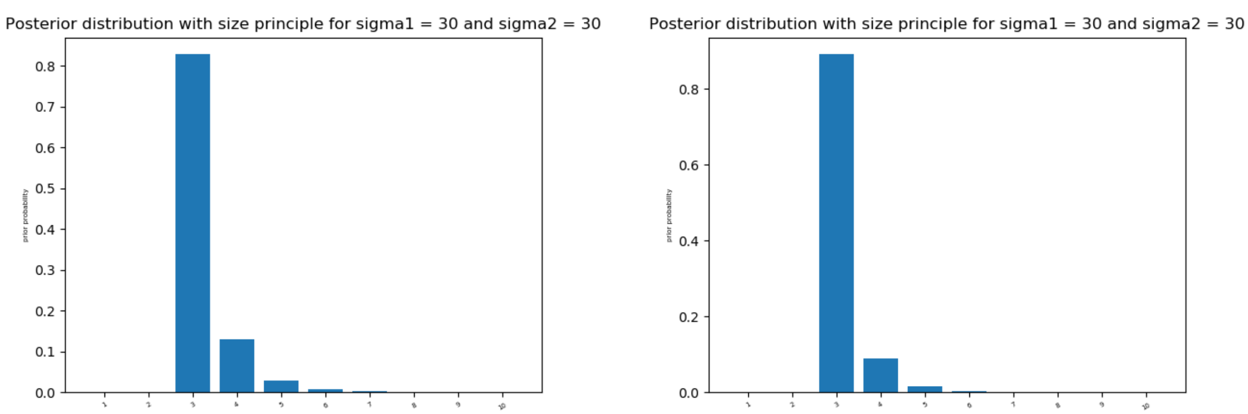
(a) expected size

(b) Uniform Prior

Figure 4: Prior probability graph : uniform and expected size prior

Hypothesis 2: Posterior probabilties for 1-3 samples $(2.2, -.2), (.5, .5), (1.5, 1)$ observed in order:

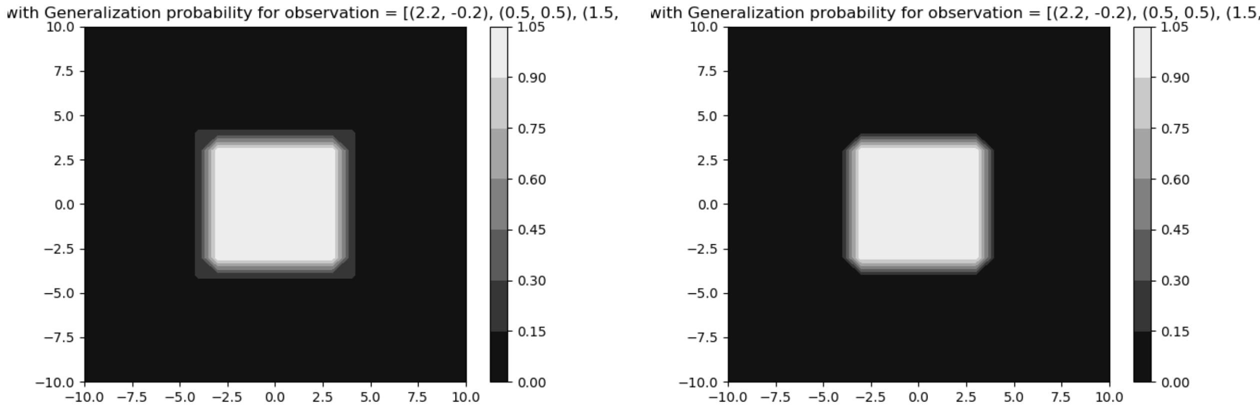
```
Uniform Prior       :  0.63 | 0.799 | 0.89
Expected-size Prior :  0.45 |  0.68 | 0.82
```



(a) expected size-posterior for 3 samples

(b) Uniform Prior-posterior for 3 samples

Figure 5: Posterior probability for $X = (2.2, -.2), (.5, .5), (1.5, 1)$ uniform and expected size prior



(a) Generalization with expected size for 3 samples

(b) Generalization with Uniform prior 3 samples

Figure 6: Generalization with expected size and uniform prior with $X = (2.2, -.2), (.5, .5), (1.5, 1)$