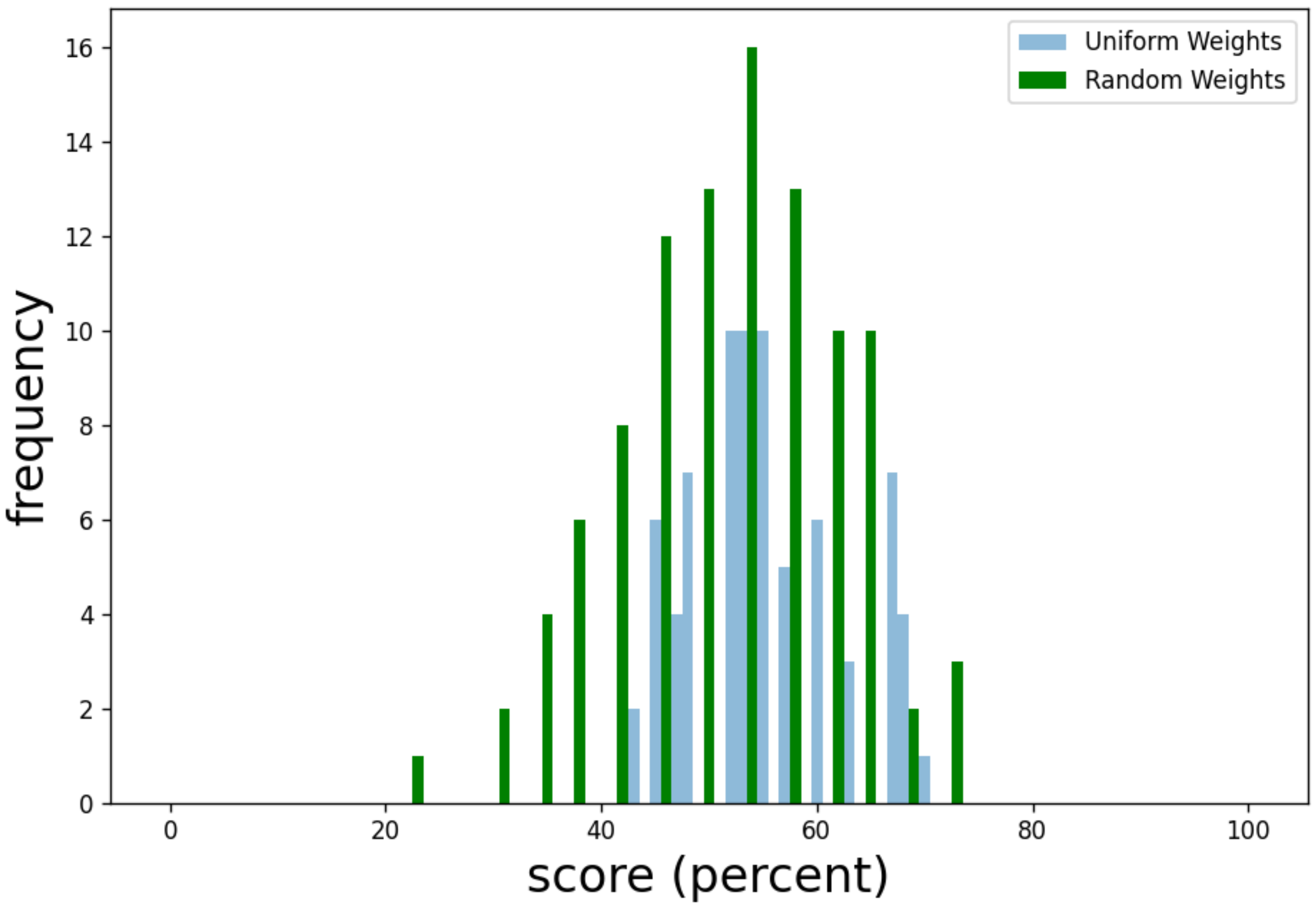
Optimal Sample Weighting

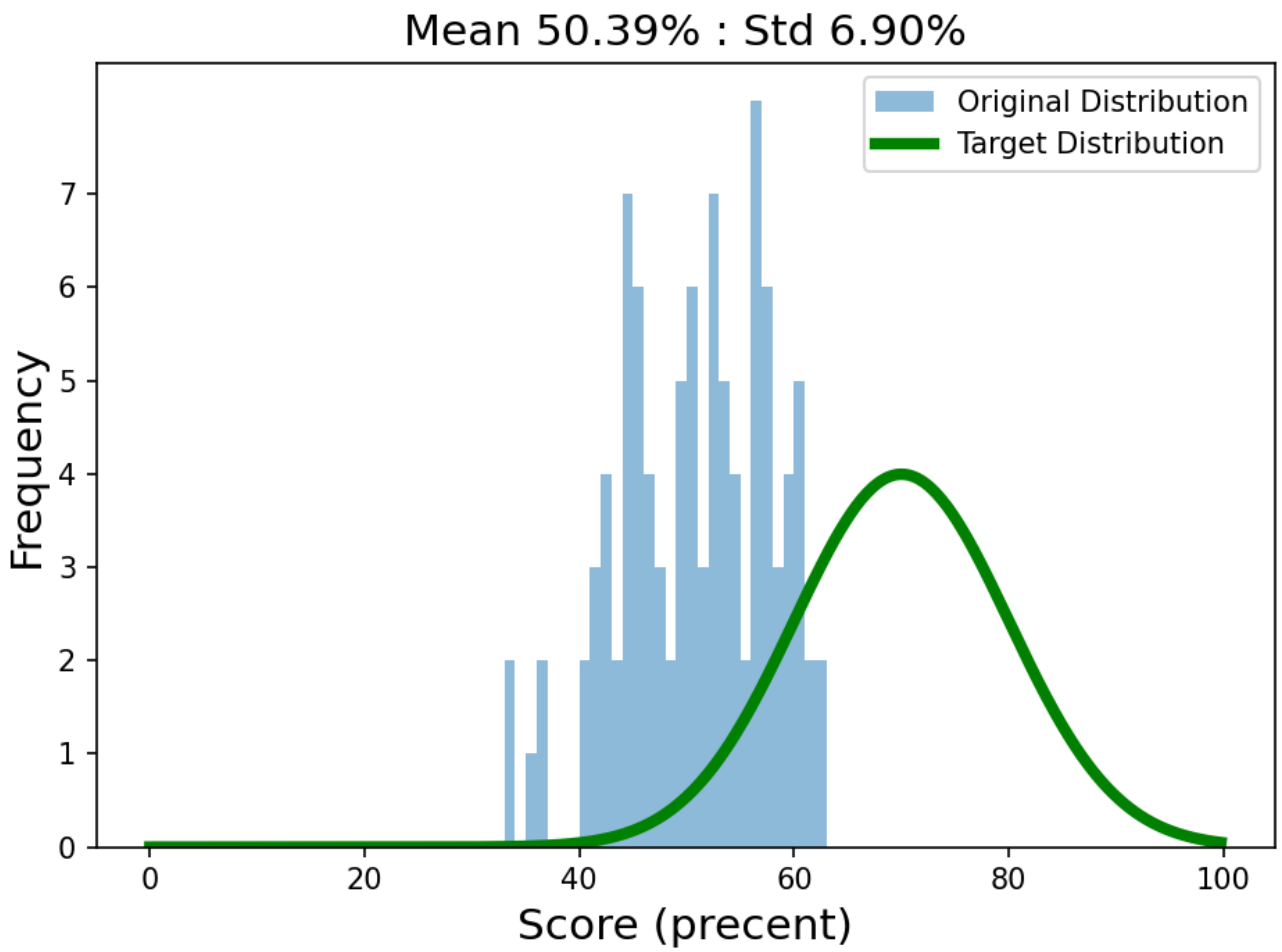
James Chok & Geoff Vasil - June 2023

Student	Q1	Q2	...	Q50
Scrooge	❌	✅		✅
Fezziwig	✅	✅		✅
...				



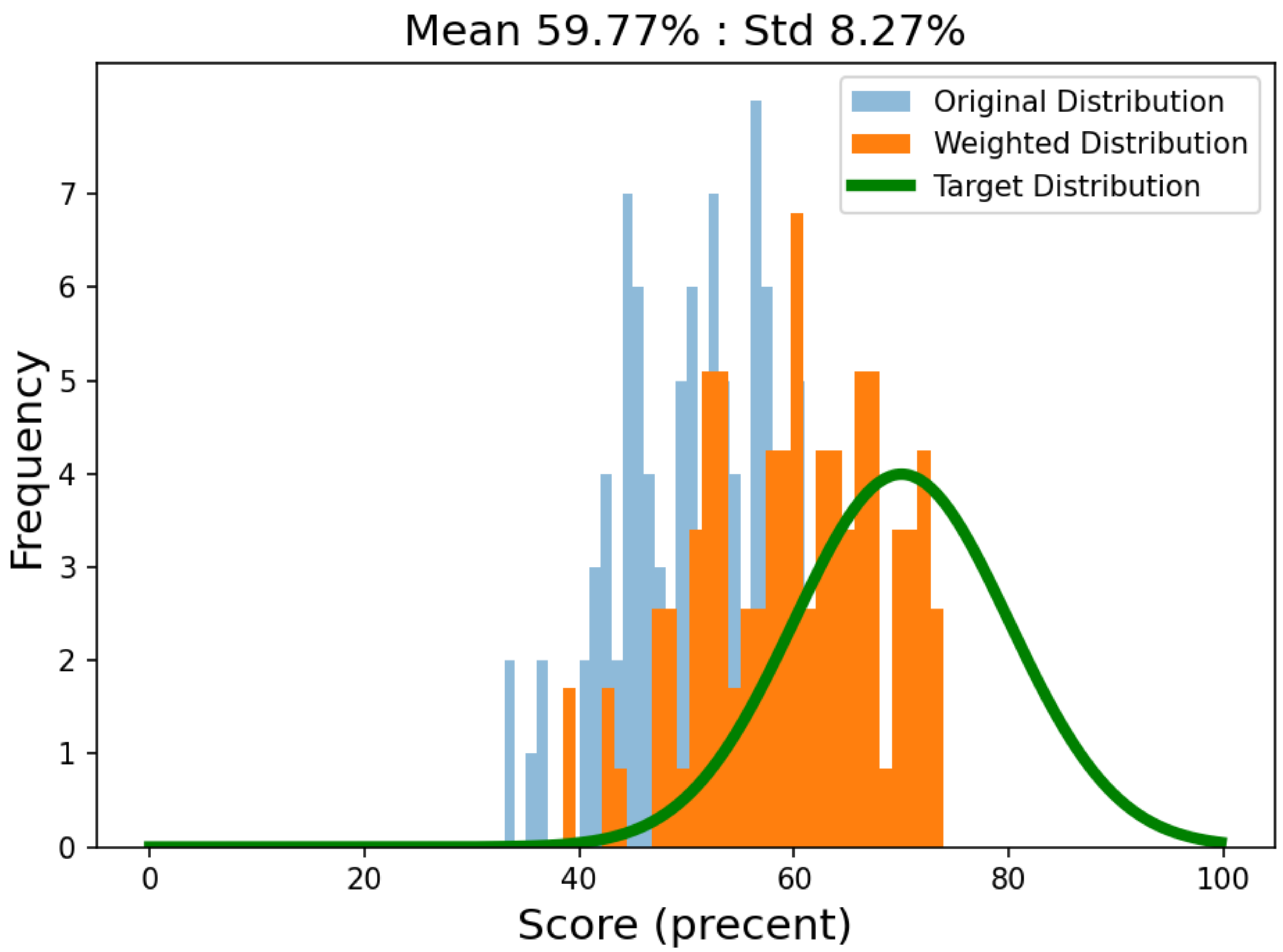
$$\text{Scrooge's Mark} = \sum_i w_i \cdot \text{(ith question is correct)}$$

Iteration 0



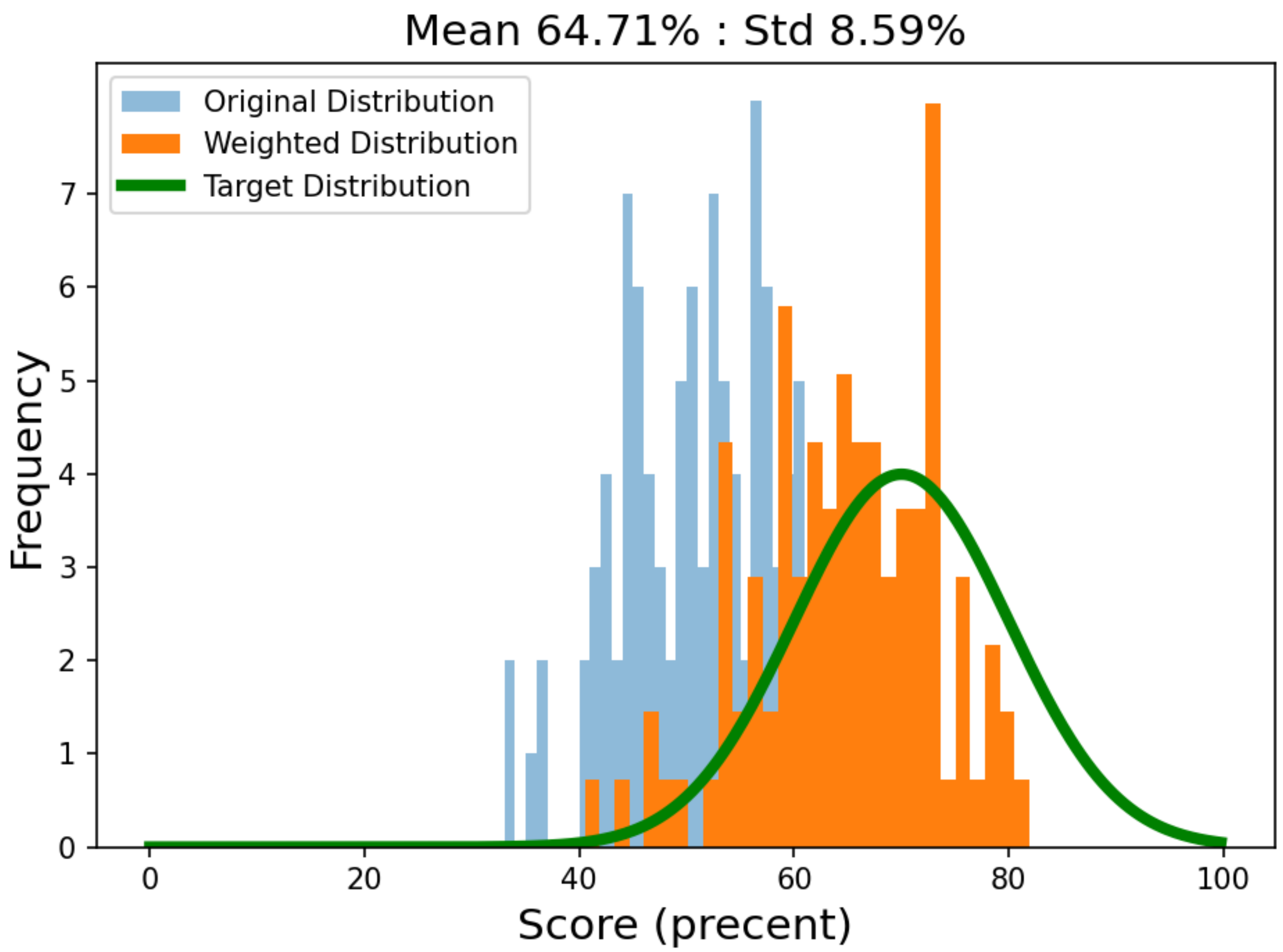
Mean 50.39% : Std 6.90%
Original Distribution
Target Distribution
Frequency
0
1
2
3
4
5
6
7
0
20
40
60
80
100
Score (precent)

Iteration 1

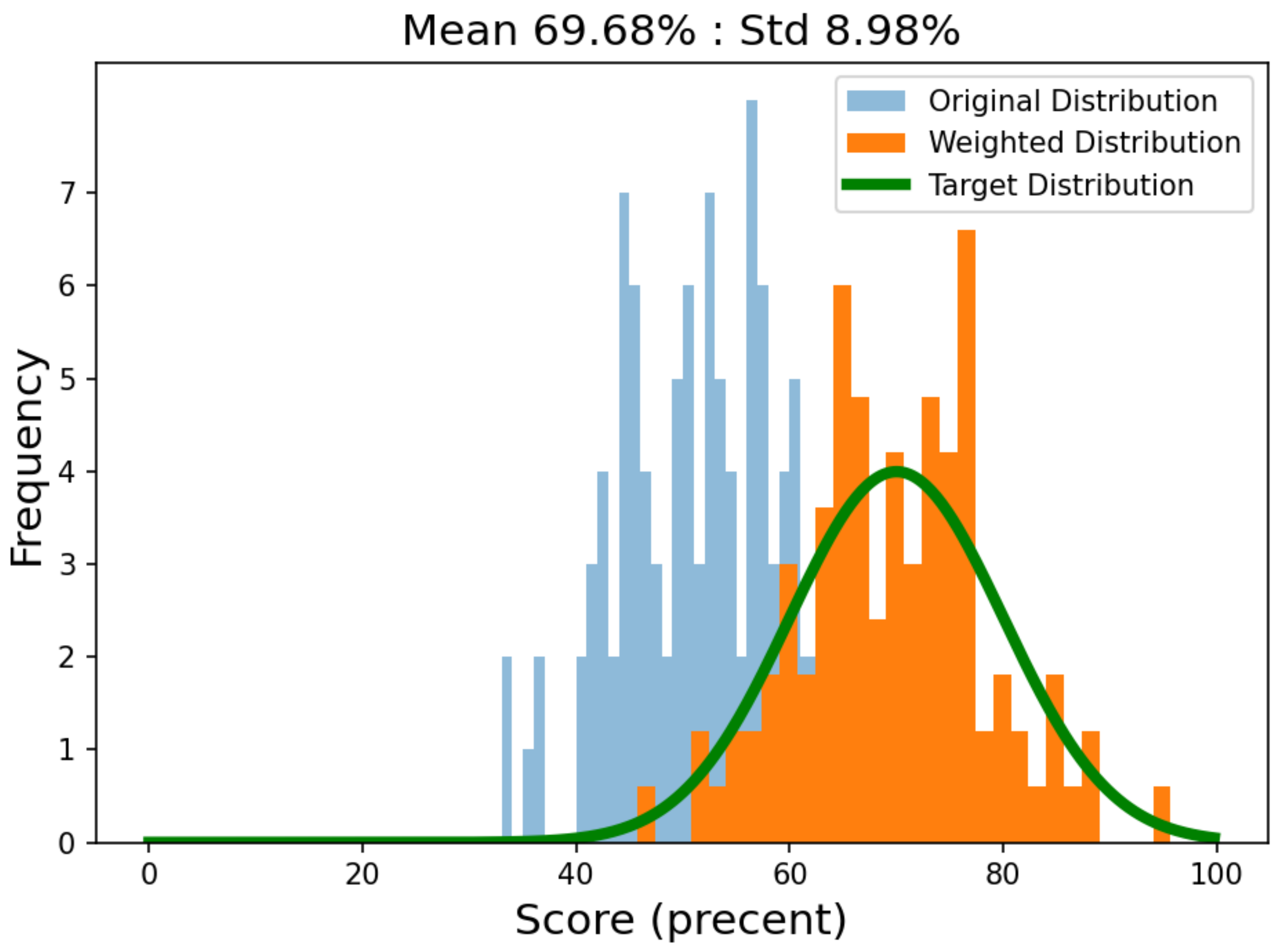


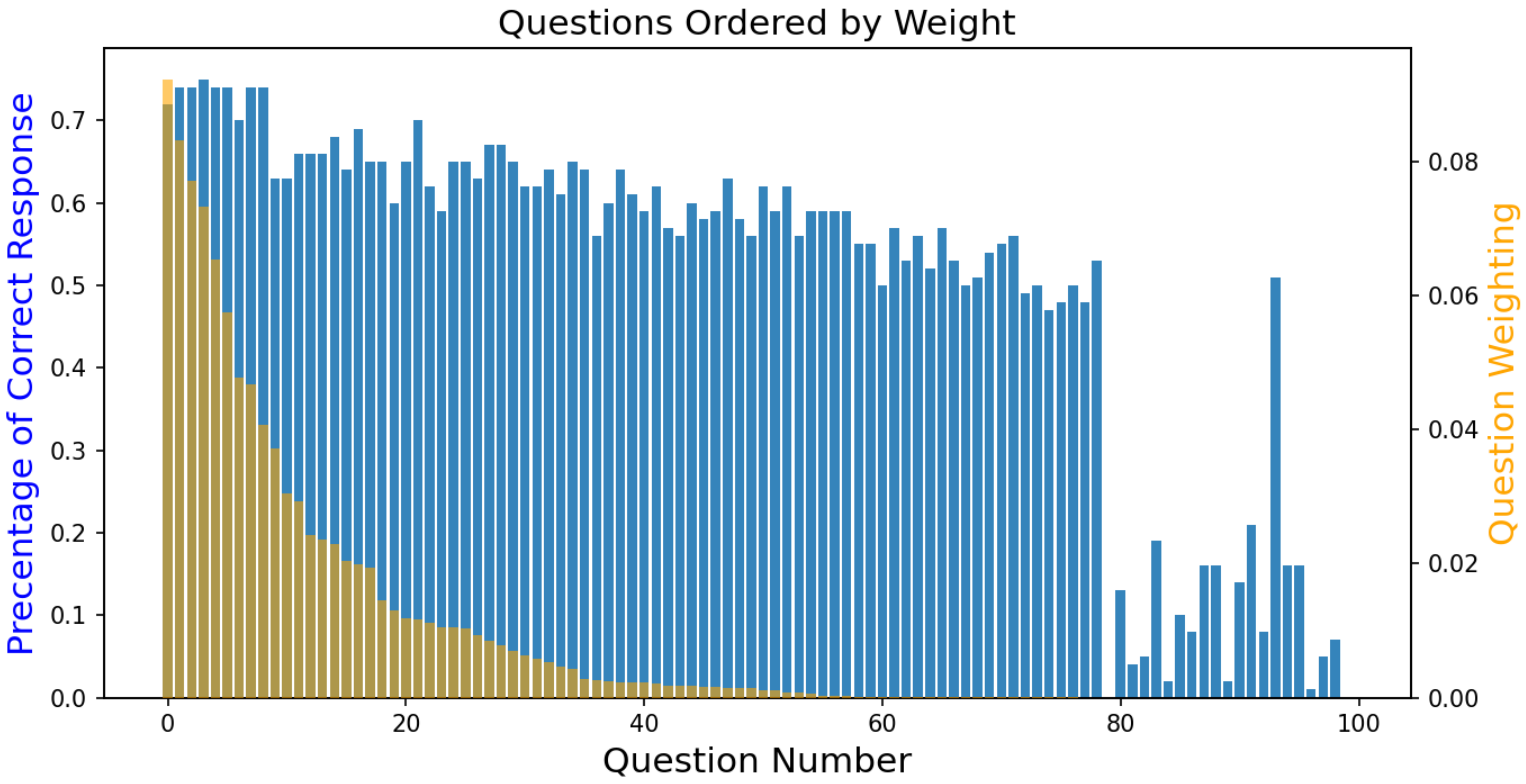
Mean 59.77% : Std 8.27%
Original Distribution
Weighted Distribution
Target Distribution
Frequency
Score (precent)
0
1
2
3
4
5
6
7
0
20
40
60
80
100

Iteration 5


Mean 64.71% : Std 8.59%
Original Distribution
Weighted Distribution
Target Distribution
Frequency
Score (precent)
0
1
2
3
4
5
6
7
0
20
40
60
80
100

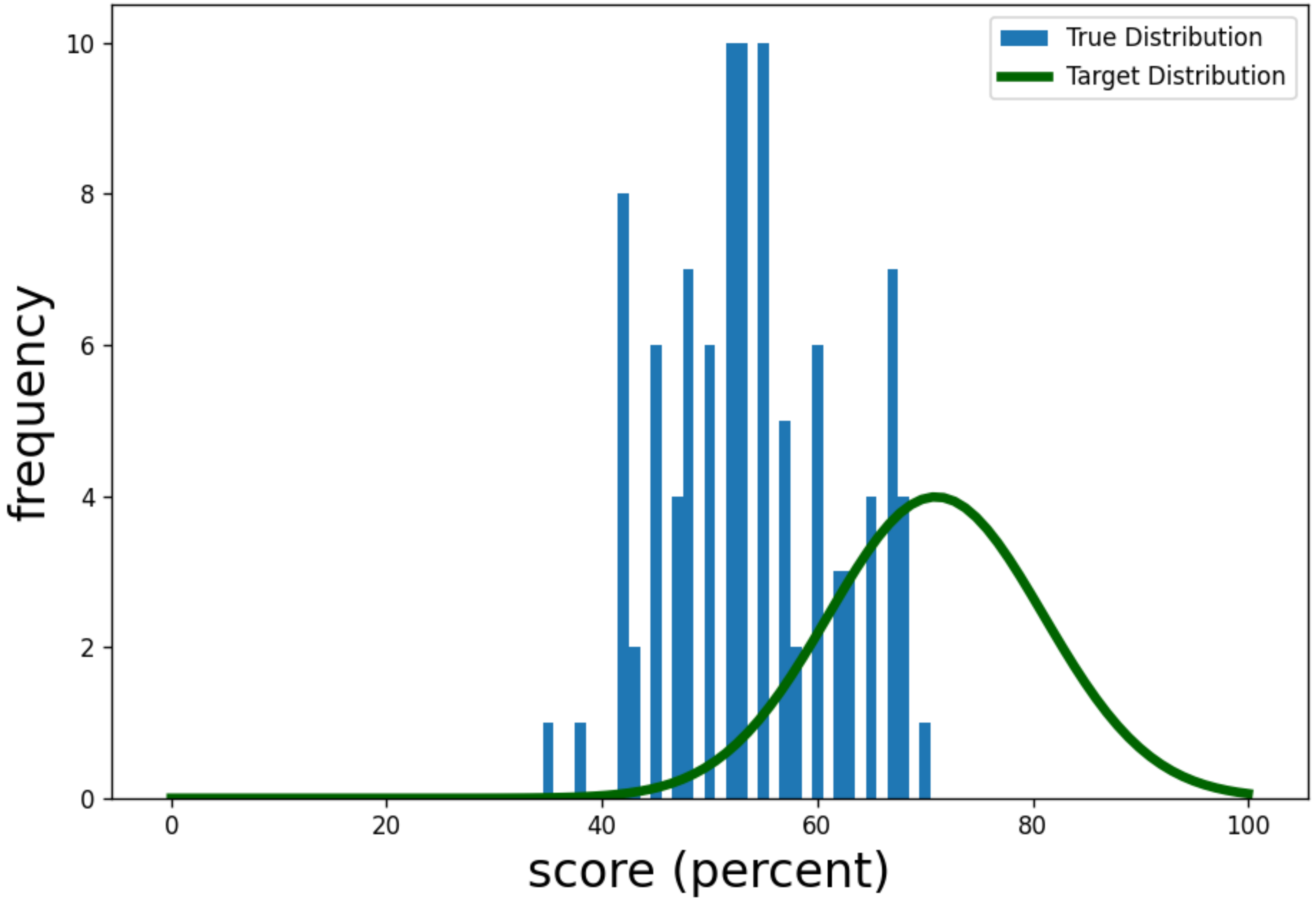
Iteration 20




Questions Ordered by Weight
Precentage of Correct Response
Question Weighting
Question Number
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.00
0.02
0.04
0.06
0.08
0
20
40
60
80
100

$$\min_{w} \; f(w)$$

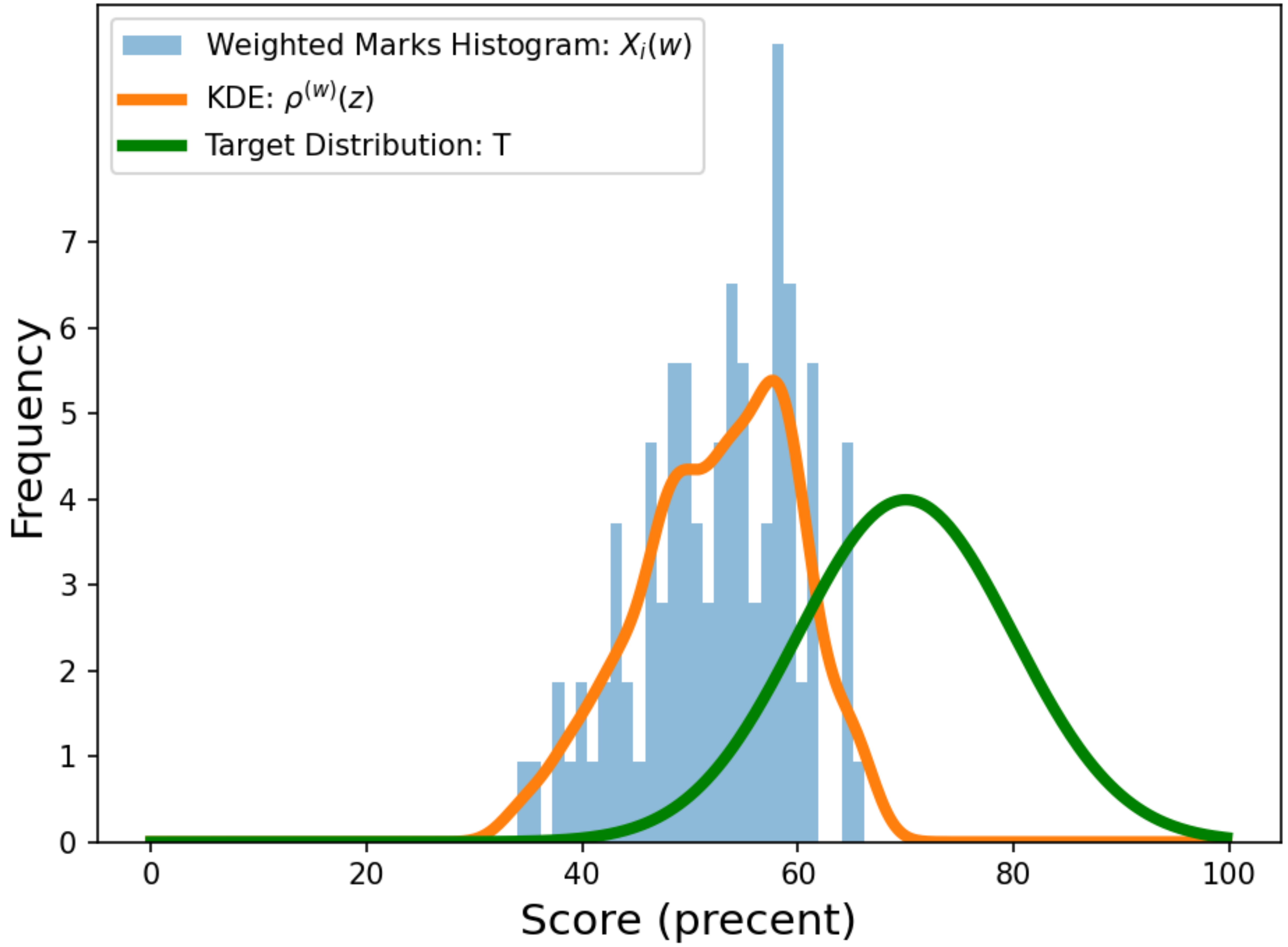
$$w_i \;\geq\; 0, \qquad \sum_i w_i = 1$$



$X_i(w)$ - Weighted mark of student i

Discrete Distribution of $X_i(w)$ $\approx$ Continuous Distribution: $\rho^{(w)}(z)$

KDE: $\rho^{(w)}(z) = \frac{1}{N}\sum_{i=1}^{N} \mu(z - X_i(w))$ $\quad$ μ is a Gaussian Kernel



$$\mathrm{KL}(\rho^{(w)}, T) = \int \rho^{(w)}(z) \log\left(\frac{\rho^{(w)}(z)}{T(z)}\right) dz$$

$$\boxed{\begin{gathered}\min_{w} \ \mathrm{KL}(\rho^{(w)}, T) \\ w_i \ \geq \ 0, \qquad \sum_i w_i = 1\end{gathered}}$$

Gradient Descent

$$w^{t+1} = w^t - \eta_t \nabla_w f$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Our Proposal

$$w^{t+1} = w^t - \eta_t w^t (\nabla_w f - w^t \cdot \nabla_w f)$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Convex optimization over a probability simplex
– James Chok and Geoffrey M. Vasil

https://arxiv.org/abs/2305.09046