A Logistic Model for Distribution of Grades on a Math Placement Exam

1 Computational Algorithm

Given a subset of students \( n \leq N \) taking a subset of questions on the MPE (which yields \( \Delta \)), the mean \( \mu \) and standard deviation \( \sigma \) of the placement test are computed. [The probabilities \( P_F \) (global fail rate) and \( P_P \) (global pass rate) are fixed throughout the following.] \( \Delta = \mu_P - \mu_F \) is the difference between passing and failing scores on the placement test (from historical data).

The algorithm proceeds as follows: Given \( n, \mu, \sigma, \Delta \)

1. The population parameters are estimated as follows
   (a) \( \mu_F = \mu - P_F \Delta \)
   (b) \( \mu_P = \mu + P_P \Delta \)
   (c) \( \alpha = 0.05 \) (empirical constant)
   (d) \( \sigma_P = \sqrt{\frac{\sigma^2 - P_P P_F \Delta^2}{P_P + P_F (1 + \alpha)^2}} \)
   (e) \( \sigma_F = (1 + \alpha) \cdot \sigma_P \)
   (f) \( N_P = P_P N \)
   (g) \( N_F = P_F N \)

2. The Logistic parameters are estimated by
   (a) \( b_F = \frac{\pi}{\sqrt{3} \sigma_P}, a_F = b_F, \mu_F \)
   (b) \( b_P = \frac{\pi}{\sqrt{3} \sigma_P}, a_P = b_P, \mu_P \)

3. The Logistic CDFs are then given by
   (a) \( \Pi_F = \Pi(a_F, b_F, S) = \frac{1}{1 + e^{a_F - b_F S}} \)
   (b) \( \Pi_P = \Pi(a_P, b_P, S) = \frac{1}{1 + e^{a_P - b_P S}} \)

4. Self Consistency. Cronbach’s \( \alpha = 0.856 \) for the 2007 data set.

\[
\begin{array}{cccccccccc}
N & \text{MPE-\mu} & \text{MPE-sigma} & \text{Date} & \text{global pass} & \text{global fail} & \text{mu} & \text{sigma} & \text{mu}_P & \text{sigma}_P \\
2007 \text{ Theory} & 1646 & 21.456 & 6.2402 & 56.22132 & 0.712 & 0.258 & 17.45104 & 5.866972 & 23.67517 & 5.614972 \\
2007 \text{ Actual} & & & & & & & 17.5112 & 5.767 & 23.4333 & 5.6393 \\
\text{error} & & & & & & & 0.33\% & 0.25\% & 0.43\% & \\
\text{error} & & & & & & & 0.32\% & 0.27\% & 2.88\% & \\
\end{array}
\]

Figure 1: Verification of Model