Can Badminton Skill Be Put on a Probability Scale?

1. From a Multivariate Ability Hypothesis to a Competitive Scale

Badminton performance depends jointly on physical condition, technical quality, tactical choice, anticipation, and consistency. The model compresses their long-run effect on match outcomes into a one-dimensional latent ability, then expresses that ability in units with an explicit win-probability interpretation.

1.1 Multivariate ability hypothesis

For player i, let X_i denote a multidimensional ability state—for example fitness, strength, speed, consistency, mentality, and tactics—and let $S_i = f(X_i)$ denote the aggregate ability relevant to long-run outcomes. Within a reasonably homogeneous training population, suppose the component effects are approximately additive on a suitable scale, only moderately dependent, and not dominated by any single factor. The central limit theorem then motivates

$$S \approx N(\mu_S, \sigma_S^2).$$

Equal aggregate ability may therefore arise from very different combinations of underlying strengths.

1.2 A level unit defined by match probability

Let L_i be player i's competitive level and $\Delta L = L_A - L_B$ the level gap. Approximate a game by 21 independent rallies and define $\Delta L = 1$ to mean that the stronger player wins the game with probability $2/3$. The binomial model then gives a rally-win probability of about 54.6% for a one-level advantage. Rally odds are used to construct an additive scale; details appear in the appendix.

Because L is a linear transformation of aggregate ability, it remains approximately normal. Let r be the fraction of the population at or above a given level, $L(r)$ the corresponding level, and Φ the standard normal CDF. Setting the median stable young male recreational player to 0 and using an empirical scale of 6.8 levels per standard deviation gives

$$\boxed{L(r) = 6.8\,\Phi^{-1}(1-r) - 15}.$$

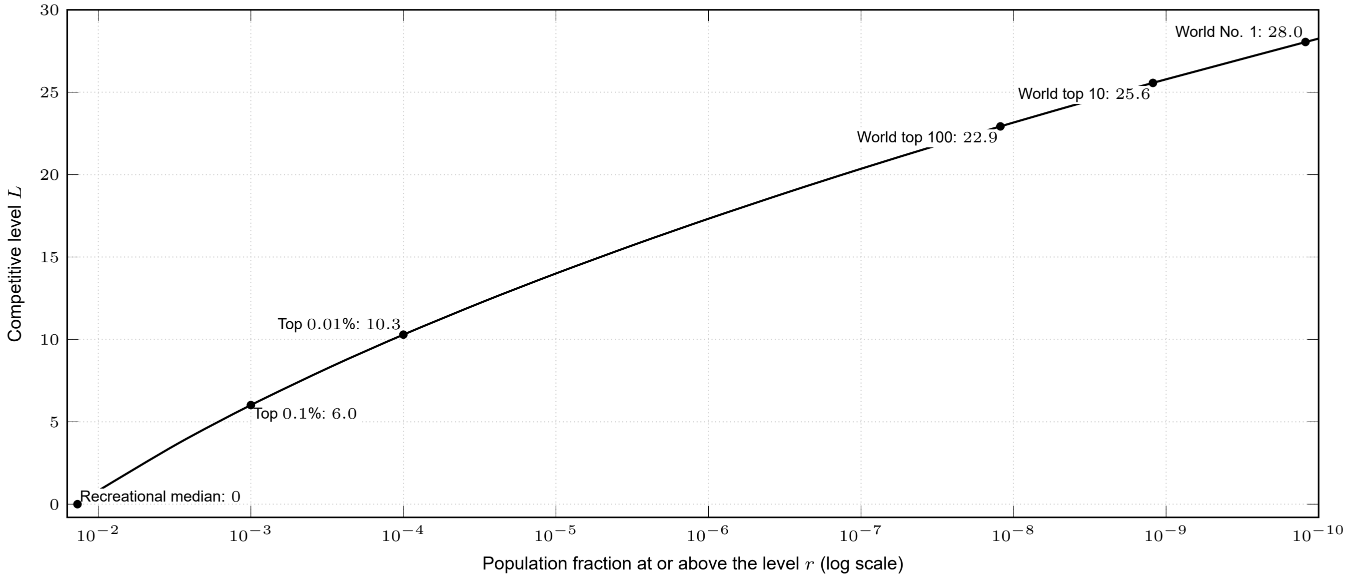
What a Level Gap Means in Competition

Level gap	Rally win	Game win	Best-of-three win	Typical score
1	54.6%	66.7%	74.1%	21:17.4
2	59.2%	80.5%	90.1%	21:14.5
4	67.8%	95.6%	99.43%	21:10.0
8	81.6%	99.95%	99.99993%	21:4.7

All values use the simplified model of a fixed 21-rally game.

2. Population-Tail Frequency and Competitive Level

Level 0 is the median stable young male recreational player. World-rank references use a total population of approximately 8.2×10^9.



3. Qualitative Checks Against Competitive Experience

Ability gaps shrink near the top. On a logarithmic population scale, the normal-quantile curve gradually flattens: each further tenfold increase in rarity corresponds to a smaller level gain. As long-run ability gaps contract, short-run fluctuations matter relatively more.

Close scores can coexist with stable outcomes. Elite players may differ only slightly in rally-win probability, yet the advantage accumulates over games and repeated matches into a reliable competitive ordering.

Rank gaps can greatly exceed ability gaps. The population density is very low in the extreme tail, so a small increase in level can correspond to a tens- or hundreds-fold change in rank.

Elite performance need not share one template. The scale constrains only the aggregate outcome of multidimensional ability. Strength, speed, control, anticipation, and other factors can compensate for one another.

These checks establish qualitative consistency, not empirical validation; the distributional assumptions and numerical calibration remain to be tested against match data.

Appendix: Formulae and Calculations

A. Multivariate Ability and the Normal Approximation

For player i, write the multidimensional ability state as $X_i = (X_{i1}, \ldots, X_{id})$ and aggregate ability as $S_i = f(X_i)$. Around a reasonably homogeneous population, take a first-order approximation:

$$S_i \approx c + \sum_{k=1}^{d} w_k X_{ik} + \varepsilon_i,$$

where ε_i collects smaller higher-order interactions. If the component contributions have finite variance, are weakly dependent, and no single term dominates, the central limit theorem motivates

$$S \approx N(\mu_S, \sigma_S^2).$$

B. An Additive Level Unit Defined by Game-Win Probability

Let p be the stronger player's probability of winning a rally. In the simplified fixed 21-rally game, the game-win probability is

$$q(p) = \sum_{j=11}^{21} \binom{21}{j} p^j (1-p)^{21-j}.$$

Define a one-level gap by $q(p_1) = 2/3$. Numerically,

$$p_1 \approx 0.54634, \qquad \kappa := \ln \frac{p_1}{1-p_1} \approx 0.1859.$$

To make level gaps additive, define them through rally odds:

$$\Delta L := \frac{1}{\kappa} \ln \frac{p}{1-p}, \qquad p(\Delta L) = \frac{1}{1+e^{-\kappa \Delta L}}.$$

The corresponding typical score is approximated by

$$21{:}x, \qquad x(\Delta L) \approx 21 \frac{1-p}{p} = 21 e^{-\kappa \Delta L}.$$

If the game-win probability is q, the probability of winning a best-of-three match is

$$P_{\text{Bo3}} = q^2 + 2q^2(1-q) = 3q^2 - 2q^3.$$

C. From a Normal Distribution to the Population-Level Curve

If competitive level L is a linear transformation of aggregate ability S, then L remains approximately normal. For $r = P(L' \geq L)$, the fraction of the population at or above level L, and standard normal CDF Φ,

$$L_{\text{abs}}(r) = \sigma_L \Phi^{-1}(1-r).$$

Take the empirical scale $\sigma_L \approx 6.8$ and place the median stable young male recreational player at level 15 on this absolute scale. The translated scale used in the main text is therefore

$$\boxed{L(r) = 6.8\Phi^{-1}(1-r) - 15}.$$

Its zero point corresponds to

$$r_0 = 1 - \Phi(15/6.8) \approx 0.01370.$$

D. World-Rank References

Using a total population $N \approx 8.2 \times 10^9$, the tail fraction for world rank m is $r = m/N$. Substitution into the curve gives:

Reference	Population-tail fraction r	Level L
World top 100	1.22×10^{-8}	22.9
World top 10	1.22×10^{-9}	25.6
World No. 1	1.22×10^{-10}	28.0