

先验分布

为每位运动员的投篮命中概率 θ_i 设定了独立的 Beta(2, 2) 先验分布。

根据要求生成的模拟数据

运动员 1 的投篮次数 $N=50$ ，真实 $\theta=0.1$ （模拟数据观测比例为 0.120）；

运动员 2 的 $N=56$ ，真实 $\theta=0.4$ （观测比例为 0.375）；

运动员 3 的 $N=68$ ，真实 $\theta=0.7$ （观测比例为 0.735）

参数设置

参数 $(\theta_1, \theta_2, \theta_3)$ 的后验分布通过马尔可夫链蒙特卡洛（MCMC）方法进行近似估计，该过程通过 R 语言的 rjags 包调用 JAGS 实现。每个 θ 运行了三个平行的 MCMC 链，每条链包含 1000 次适应性迭代和 10000 次 burn-in。预烧之后，每条链运行 20000 次迭代，并设置抽样间隔（thinning）为 2，最终每条链保存 10000 个样本，总计 30000 个样本用于后验分析。

为了评估差异是否具有实际意义，定义了一个实际等效区域（Region of Practical Equivalence, ROPE），范围为 $[-0.05, 0.05]$ 。

MCMC 收敛性评估

MCMC 链的收敛性对于获得可靠的后验估计至关重要。通过 Gelman-Rubin Diagnostic 计算了每个 θ 参数的潜在尺度缩减因子（PSRF，或 R-hat）。结果显示，所有参数的点估计值均为 1.00，其置信区间上限也为 1.00。接近 1.0 的值表明链已成功收敛，即不同的链收敛到了相同的目标分布。有效样本量（Effective Sample Size, ESS）的值在 9591 至 9999 之间，表明自相关性低，后验分布的抽样效率高。基于以上诊断结果，可以认为 MCMC 模拟已成功收敛，所得样本为后验分布提供了可靠的近似。

各运动员后验分布Summary

关于个体运动员的命中概率 θ ，后验分布结果显示：

运动员 1 (θ_1) 的后验均值为 0.148（中位数 0.144，众数 0.130），其 95% 最高密度区间 (HDI) 为 $[0.059, 0.240]$ ，此区间代表了运动员 1 投篮命中概率最可信的取值范围。

运动员 2 (θ_2) 的后验均值为 0.383 (中位数 0.381, 众数 0.374) , 95% HDI 为 [0.266, 0.509]。

运动员 3 (θ_3) 的后验均值为 0.722 (中位数 0.724, 众数 0.727) , 95% HDI 为 [0.622, 0.828]。

运动员之间

关于命中概率的成对差异 ($\theta_i - \theta_j$) , 后验分析表明: 差异 ($\theta_1 - \theta_2$) 的后验均值为 -0.235, 95% HDI 为 [-0.389, -0.082]。由于该区间完全不包含 0 且和 ROPE 无交集, 存在极强的证据表明 θ_1 可信地低于 θ_2 。

差异 ($\theta_1 - \theta_3$) 的后验均值为 -0.574, 95% HDI 为 [-0.708, -0.428]。该区间同样完全为负值且和 ROPE 无交集, 提供了极强的证据表明 θ_1 可信地低于 θ_3 。

差异 ($\theta_2 - \theta_3$) 的后验均值为 -0.340, 95% HDI 为 [-0.495, -0.178]。该区间的上限为负数且区间和 ROPE 无交集, 提供了极强的证据表明 θ_2 可信地低于 θ_3 。

在 ROPE 分析方面, 使用 [-0.05, 0.05] 作为实际等效区域来判断差异是否在实践中可忽略。对于差异 ($\theta_1 - \theta_2$) , 其 95% HDI [-0.389, -0.082] 完全位于 ROPE 之下, 仅有 1.03% 的后验概率质量落在 ROPE 内, 而 98.95% 落在其下方, 这强烈表明该差异是可信的。对于差异 ($\theta_1 - \theta_3$) , 其 95% HDI [-0.708, -0.428] 完全位于 ROPE 之下, 几乎 100% (四舍五入后) 的后验概率质量落在 ROPE 之下, 这提供了极强的证据表明该差异是可信的。对于差异 ($\theta_2 - \theta_3$) , 其 95% HDI [-0.495, -0.178] 也完全位于 ROPE 之下, 仅有 0.03% 的后验概率质量落在 ROPE 内, 而 99.97% 落在其下方, 同样提供了极强的证据表明该差异是可信的。