

表 1. 变量分布趋势及 Pearson 相关系数表

	M	SD	Win	ERA	BB	RBI	Slug
Win	49.98	7.53	--				
ERA	3.86	0.4	-0.746***	--			
BB	488	57.52	0.520**	-0.256	--		
RBI	642.37	70.79	0.725***	-0.225	0.529**	--	
Slug	0.4	0.02	0.601***	-0.084	0.383*	0.903***	--

Note: *: p < 0.05; **: p < 0.01; ***: p < 0.001.

Pearson 相关系数分析结果如表 1 所示：球队胜场数 (Win) 与投手防御率 (ERA) 之间存在显著负相关关系 ($r=-0.746$, $p<0.001$)，说明投手失分能力越强，球队获胜场数越多，反映了投手表现对比赛结果的关键影响；Win 与进攻类指标 BB、RBI 及 Slug 均呈现显著正相关，其中 Win 与 RBI 的相关系数最高 ($r=0.725$, $p<0.001$)，其次为 Slug ($r=0.601$, $p<0.001$) 和 BB ($r=0.520$, $p<0.01$)，表明球队得分能力与进攻效率对胜场数具有显著促进作用。进一步考察自变量之间的相关关系发现，ERA 与各进攻变量之间相关性较弱且均未达到显著水平，说明投手指标与进攻指标在统计意义上相对独立；然而，RBI 与 Slug 之间存在极强正相关 ($r=0.903$, $p<0.001$)，两者可能反映相似的进攻能力维度，在后续多元回归分析中需关注潜在的多重共线性问题。

表 2. 多元线性回归 (Hierarchical) 案例结果 (推荐汇报格式)

Model				Coeff.								
R-Squ	Adj. R-Squ	F(df1,df2)	Sig.	B	SE	Beta	sr^2	η_p^2	t	Sig.	VIF	
(Constant)				54.231	7.450				7.280***	<0.001		
ERA ¹	0.883	0.874	101.927(2,27)***	<0.001	-11.482	1.264	-0.614	0.358	0.754	-9.085***	<0.001	1.053
RBI ²				0.062	0.007	0.587	0.327	0.737	8.691***	<0.001	1.053	

Note: ¹: model1: (Constant), ERA: R-Squ = 0.556; ²: model2: (Constant), ERA, RBI: R-Squ = 0.883, Adj. R-Squ = 0.874, $\Delta R-Squ = 0.327$, $\Delta F(df1,df2) = 75.539(1,27)***$. ***: p < 0.001.

为探究自变量 (ERA、RBI、BB、Slug) 对因变量 (Win) 的影响, 为减少候选自变量数量并辅助模型设定, 首先采用逐步回归对候选变量进行探索性分析, 其结果不直接用于推断, 最终模型依据理论假设, 通过分层回归加以确认, 结果如表 2 所示: 模型整体拟合效果良好, $R^2=0.883$, 调整后的 $R^2=0.874$, 表明自变量共同解释因变量 87.4% 的变异; 方差分析结果表明模型具有统计学意义 ($F(2,27)=101.927$, $p<0.001$)。

在包含 ERA 与 RBI 的多元线性回归模型中, 常数项为 54.231 ($t=7.280$, $p<0.001$) ; 自变量 ERA 对因变量 (Win) 有显著负向影响 ($B=-11.482$, $\beta=-0.614$, $t=-9.085$, $p<0.001$, $\eta_p^2=0.754$, $\text{sr}^2=0.358$) , 且当控制 RBI 后, ERA 每增加 1 个标准差, 因变量 (Win) 平均减少 0.614 个标准差, ERA 的条件效应量较大, 其独特贡献解释了因变量 (Win) 总变异的 35.8%, 表明 ERA 在模型中具有较高独特性的预测作用; 自变量 RBI 对因变量 (Win) 有显著正向影响 ($B=0.062$, $\beta=0.587$, $t=8.691$, $p<0.001$, $\eta_p^2=0.737$, $\text{sr}^2=0.327$) , 且当控制 ERA 后, RBI 每增加 1 个标准差, 因变量 (Win) 平均增加 0.587 个标准差, RBI 同样具有较大的条件效应量, 其独特贡献解释了因变量 (Win) 总变异的 32.7%, 表明 RBI 亦为模型中的核心预测变量。

此外, 通过嵌套模型对比分析可知, 模型 1 的判定系数为 0.556, 模型 2 的判定系数提升至 0.883 ($\Delta R^2=0.327$) , 且增量显著 ($\Delta F(1,27)=75.5$

39, $p<0.001$)。

模型诊断结果显示, ERA 与 RBI 的方差膨胀因子均小于 5, 未发现多重共线性问题; 残差正态性检验未显示显著偏离正态分布 ($p>0.05$), 表明模型在主要诊断指标上基本满足多元线性回归的假设条件。总体而言, 基于分层回归构建的模型在当前样本中具有良好的统计适配性与解释力。

表 3. 多元线性回归 (Hierarchical) 案例结果 (推荐汇报格式)

Model				Coeff.							
R-Squ	Adj. R-Squ	F(df1,df2)	Sig.	B	SE	Beta	R^2	η_p^2	t	Sig.	VIF
(Constant)				54.231	7.450				7.280***	<0.001	
RBI ¹	0.883	0.874	101.927(2,27)***	<0.001	0.062	0.007	0.587	0.327	0.737	8.691***	<0.001
ERA ²				-11.482	1.264	-0.614	0.358	0.754	-9.085***	<0.001	1.053

Note: ¹: model1: (Constant), RBI: R-Squ = 0.525; ²: model2: (Constant), RBI,ERA: R-Squ = 0.883, Adj. R-Squ = 0.874, Δ R-Squ = 0.358, Δ F(df1,df2) = 82.544(1,27)***. ***: $p < 0.001$.

表 4. 多元线性回归 (Hierarchical) 案例结果

(条件过程分析数据集: ex11-2)

	Model 1				Model 2				Model 3			
	Beta	t	sr ²	VIF	Beta	t	sr ²	VIF	Beta	t	sr ²	VIF
(常量)		8.891***				3.188**				3.137**		
性别	0.020	0.363	< .000	1.149	0.006	0.102	< .000	1.158	0.003	0.047	< .000	1.159
婚姻	-0.011	-0.157	< .000	1.820	0.009	0.132	< .000	1.853	0.005	0.077	< .000	1.854
年龄	-0.016	-0.197	< .000	2.368	-0.011	-0.143	< .000	2.425	-0.007	-0.084	< .000	2.427
职业	0.026	0.417	< .000	1.398	0.019	0.321	< .000	1.401	0.013	0.214	< .000	1.404
教育	0.085	1.492	.006	1.183	0.073	1.334	.005	1.184	0.081	1.476	.006	1.189
月收入	-0.034	-0.440	.001	2.135	0.017	0.230	.000	2.179	0.010	0.130	< .000	2.184
消费次数	0.100	1.712	.008	1.228	0.058	1.028	.003	1.255	0.071	1.262	.004	1.268
IM					0.238	4.628***	.057	1.036	0.245	4.783***	.061	1.040
SC					0.186	3.561***	.035	1.071	0.179	3.444***	.032	1.074
IM_mcxSC_mc									-0.116	-2.277*	.014	1.027
F(df1,df2)	0.804(7,356)				4.161(9,354)***				4.307(10,353)***			
R ²	0.016				0.096				0.109			
Adj. R ²	-0.004				0.073				0.084			
ΔR ²	0.016				0.080				0.013			
ΔF(df1,df2)	0.585(7,356)				15.678(2,354) ***				5.184(1,353) *			