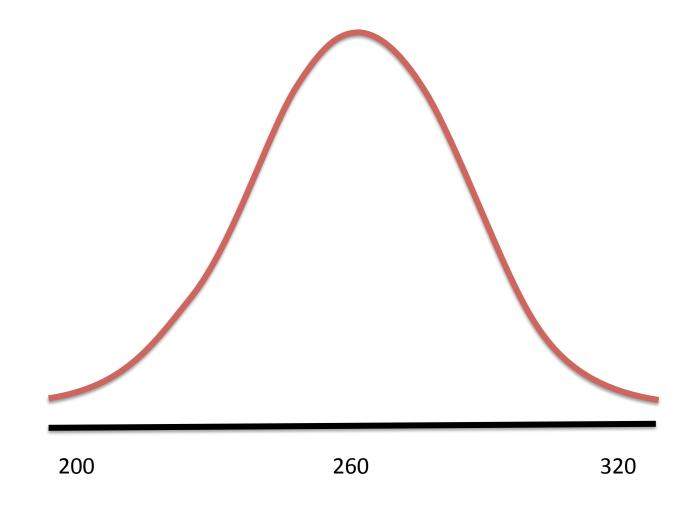


DATE	OPP	RESULT	AB	R	н	2B	3B	HR	RBI	ВВ	SO	SB	CS	OBP	SLG	OPS	AVG
Apr 1	@ NYY	W 8-2	5	1	3	0	0	0	1	0	1	0	0	.600	.600	1.200	.600
Apr 3	@ NYY	W 7-4	4	1	2	1	0	0	0	0	1	0	0	.556	.667	1.223	.556
Apr 4	@ NYY	L 4-2	3	0	2	0	0	0	0	0	0	0	0	.583	.667	1.250	.583
Apr 5	@ TOR	W 6-4	0	0	0	0	0	0	0	0	0	0	0	.615	.667	1.282	.583
Apr 6	@ TOR	L 5-0	Did not play														
Apr 7	@ TOR	W 13-0	5	1	2	1	0	0	0	0	1	0	0	.556	.647	1.203	.529
Apr 8	vs BAL	W 3-1	3	0	0	0	0	0	0	0	0	0	0	.476	.550	1.026	.450
Monthly Totals			20	3	9	2	0	0	1	0	3	0	0	.476	.550	1.026	.450



A (rough) sketch of the MLB batting average distribution.

A hierarchical model

$$heta \sim N(\mu, \tau^2)$$

$$Y|\theta \sim N(\theta, \sigma^2)$$

Here, θ denotes *any* batting average among the MLB players, and Y denotes the player's batting average. The parameter τ quantifies the prior standard deviation, and σ describes the sampling standard deviation. Specifically:

$$\theta \sim N(260, 34^2)$$

$$Y|\theta \sim N(\theta, 110^2)$$

A hierarchical model

Best guess for the players batting average, given the observed data:

$$E(\theta|Y) = B\mu + (1 - B)Y$$
$$= \mu + (1 - B)(Y - \mu)$$
$$B = \frac{\sigma^2}{\sigma^2 + \tau^2}$$

Specifically:

$$E(\theta|Y = 450) = B \times 260 + (1 - B) \times 450$$
$$= 260 + (1 - B)(450 - 260)$$
$$B = \frac{110^2}{110^2 + 34^2}$$
$$E(\theta|Y = 450) \approx 270$$