Lecture 20: Modeling Rare Events

Instructor: Ryan Brill *Scribe: Jonathan Pipping*

20.1 NBA Box Scores

20.1.1 Lebron's Career Average

Lebron James has played about 1,800 games in his NBA career. Over his career, he's averaged 27 points, 7 rebounds, and 7 assists per game, but has never had a game with this exact box score! How unlikely is this? We'll start with a rough estimate of the probability this occurs in a single game. First, let's make the (false) assumption that the number of points, rebounds, and assists are independent. Then we assign some probability to each of these events.

$$\begin{aligned} \mathbb{P}(27 \text{ points}) &= 1/20 \\ \mathbb{P}(7 \text{ rebounds}) &= 1/10 \\ \mathbb{P}(7 \text{ assists}) &= 1/10 \end{aligned}$$

Then since we assumed independence, we have the following probability for this box score:

$$\begin{aligned} \mathbb{P}(27 \text{ points}, 7 \text{ rebounds}, 7 \text{ assists}) &= \mathbb{P}(27 \text{ points})\mathbb{P}(7 \text{ rebounds})\mathbb{P}(7 \text{ assists}) \\ &= \frac{1}{20} \times \frac{1}{10} \times \frac{1}{10} = \frac{1}{2000} \end{aligned}$$

This is a rough estimate, but the probability looks pretty small! Clearly, realizing this box score in a single game is a rare event.

20.1.2 Realizing a Rare Event

Now that we have a rough estimate of the probability of this box score occurring in a single game, how about the probability that this box score doesn't occur in **any** of n games? We set up a probability model to answer this question. Let X be the number of games in which Lebron James scores 27 points, gets 7 rebounds, and gets 7 assists. Then X is a binomial random variable with parameters $n = 1800$ and $p = 1/2000$.

$$X \sim \text{Binomial}(n = 1800, p = \frac{1}{2000})$$

Then the probability that Lebron James never scores 27 points, gets 7 rebounds, and gets 7 assists in n games is given by:

$$\begin{aligned} \mathbb{P}(X = 0) &= \binom{1800}{0} \left(\frac{1}{2000}\right)^0 \left(\frac{1999}{2000}\right)^{1800} \\ &= \left(\frac{1999}{2000}\right)^{1800} \\ &\approx 0.406 \end{aligned}$$

Still, this computation is a bit tedious, and we'd like a cleaner way to estimate the probability of this rare event. This leads us to the Law of Rare Events, or the Poisson Limit Theorem.

20.1.3 Poisson Limit Theorem

Theorem 20.1 (Poisson Limit Theorem). *Suppose $X \sim$ Binomial(n, p_n), where $\lim_{n\to\infty} np_n = \lambda \in (0, \infty)$. Then $X \xrightarrow{d}$ Poisson(λ).*

Proof. We begin with the probability mass function of the binomial distribution:

$$\begin{aligned}\mathbb{P}(X = k) &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= \frac{n!}{k!(n-k)!} p_n^k (1 - p_n)^{n-k}\end{aligned}$$

Now we multiply and divide by n^k to get the following:

$$\begin{aligned}\mathbb{P}(X = k) &= \frac{n!}{k!(n-k)!} \frac{(np_n)^k}{n^k} \left(1 - \frac{np_n}{n}\right)^{n-k} \\ &= \frac{n!}{n^k(n-k)!} \frac{(np_n)^k}{k!} \left(1 - \frac{np_n}{n}\right)^{n} \left(1 - \frac{np_n}{n}\right)^{-k}\end{aligned}$$

Now we take the limit as $n \to \infty$. Note the following:

$$\begin{aligned}\frac{n!}{n^k(n-k)!} &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \cdot \frac{(n-k)!}{(n-k)!} \\ &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \xrightarrow{n} 1\end{aligned}$$

$$\frac{(np_n)^k}{k!} \approx \frac{\lambda^k}{k!}$$

$$\left(1 - \frac{np_n}{n}\right)^{n} \approx \left(1 - \frac{\lambda}{n}\right)^{n} \xrightarrow{n} e^{-\lambda}$$

$$\left(1 - \frac{np_n}{n}\right)^{-k} \approx \left(1 - \frac{\lambda}{n}\right)^{-k} \xrightarrow{n} 1$$

So as $n \to \infty$, we have the following:

$$\begin{aligned}\mathbb{P}(X = k) &\approx \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \mathbb{P}\left(\text{Poisson}\left(\lambda\right) = k\right)\end{aligned}$$

So the proof is complete. $\square$

20.1.4 Applying the Theorem to Our Example

Let $n = 1800$ and $p = 1/2000$. Then $\lambda = np = 0.9$. Let X be the number of games in which Lebron James scores 27 points, gets 7 rebounds, and gets 7 assists. Then X is a binomial random variable with parameters

$n = 1800$ and $p = 1/2000$. So we have the following from the Poisson Limit Theorem:

$$X \xrightarrow{d} \text{Poisson}(0.9)$$

Then we can compute the probability that Lebron James never scores 27 points, gets 7 rebounds, and gets 7 assists in 1800 games as follows:

$$\begin{aligned}
\mathbb{P}(X = 0) &\approx \mathbb{P}\left(\text{Poisson}(0.9) = 0\right) \\
&= e^{-0.9} \\
&= 0.407
\end{aligned}$$

So the probabiilty that Lebron James never scored 27 points, got 7 rebounds, and got 7 assists in 1800 games is approximately 40.7%. So it's not that surprising we haven't seen him attain this box score before!