Chapter 4

Statistical Methods

4.1 Introduction

In an ideal world, we would have unlimited amount of data, and all relevant questions could be answered with certainty. Is Tom Brady better than Peyton Manning? Have them played hundreds and hundreds of games with the same teammates against similar opponents and analyze the results. Of course, in the real world, this is not possible, and we have to base our analyses on the available data.

Statistical methods play at least two roles in these situations. First, they provide methods for extracting the maximum amount of information from a dataset. Second, they give us a way to quantify the uncertainty that results from having to base these conclusion on such limited data. The goal of this chapter is to give an overview of statistical reasoning and the type of statistical methods that are useful in analyzing sports data.

4.2 Using the Margin of Error to Quantify the Variation in Sports Statistics

One goal of analytic methods is to use data to draw conclusions about the process generating data. For instance, suppose we want to study the performance of an NFL running back. Let Y denote the yards gained on a particular carry and let $Y_1, Y_2, \ldots, Y_n$ denote the results of the running back's carries in a given season.

Then, Y can be modeled as a random variable; its distribution is a way to describe the running back's "true ability level," which plays a central role in his performance. This true ability level is unknown, but its properties are reflected in the observed data. A simple model for relating the observed result to this hypothetical true ability level is to assume that $Y_1, Y_2, \ldots, Y_n$ are independent random variables each with distribution of Y. More complicated models might take into account the fact that different carries occur in different situations (e.g., down, distance to first down) and against different defenses. However, for now, the focus is on the simple model.

Under the simple model, the probability distribution of Y completely characterize the running back's ability on running plays. Therefore, if we know this distribution, we can use it to help predict future results, to compare this player to other running backs, to aid in play calling, and so on. However, because the distribution is unknown, we must use the information in $Y_1, Y_2, \ldots, Y_n$ to learn about the probability distribution of Y. The practice of using sample data to determine the properties of the underlying probability distributions is knows as statistical estimation or simply as estimation.

In many cases, estimation can be based on the analogy principle. To estimate a characteristic of the distribution of Y, we use the corresponding characteristic of the data $Y_1, Y_2, \ldots, Y_n$. For instance, to estimate the mean of Y, we use the sample mean of $Y_1, Y_2, \ldots, Y_n$; to estimate the standard deviation of Y, we use the sample standard deviation of $Y_1, Y_2, \ldots, Y_n$, and so on.

Suppose the running back in question is Jamaal Charles in the 2010 NFL season. We can estimate properties of the distribution of Y. For instance, the sample mean of the observed Y values is 6.38 yards per carry; therefore, our estimate of the mean of Y is 6.38 yards per carry. Similarly, the sample median of the observed Y values is 4 yards, yielding an estimate of the median of Y.

When interpreting results of this type, it is important to recognize the inherent variability in sports results. For instance, if we were somehow able to replay the 2010 NFL season, we would not expect Charles to have a yards-per-carry value of exactly 6.38 again, although we would expect it to be "close to 6.38." Furthermore, because of this variability, we know that the observed statistics do not exactly measure a player's true ability; for instance, we do not expect Charles's "true yards per carry" for 2010 to be exactly 6.38 yards, although, again, we expect that it will be close to 6.38 yards. Here we can think of his true yards per carry as the yards per carry that would be obtained of we were able to observe a large number of rushing attempts by Charles; alternatively, we can think of it as the mean of a random variable Y representing the results of a Charles rushing attempts.

Of course, simply saying that if we were to replay the 2010 NFL season, Charles's yards per carry would be close to 6.38 yards or his true yards per carry is close to 6.38 yards is not enough; we need a quantitative measure of what *close to* means. In both cases, this is given by a statistical measure known as the *margin of error*.

To understand what the margin of error is measuring, consider the following experiment: suppose we simulate Charles's performance during the 2010 NFL season by randomly choosing the results of 230 rushing attempts; recall that in the actual 2010 NFL season, Charles had 230 rushing attempts. An important consideration in carrying out this experiment is the probability distribution to use for the simulation. The best information we have about the distribution of the yards gained on Charles's rushing attempts is the actual results. Therefore, to simulate the result of his first rushing attempt, we randomly select one result from his 230 actual results; to simulate the result of his second rushing attempt, we select another result from his 230 actual results, and so on. Doing this 230 times gives us a simulated season for Charles; when I carried such a process, I observed a simulated yards per carry of 6.54 yards.

The difference between the simulated value of 6.54 and the actual value of 6.38 gives us important information regarding the natural variability in Charles's yards per carry. However, we could obtain more information by simulating the 2010 season several times; I did this, obtaining yards-per-carry values of 6.37, 6.29, 6.15, 5.71, 7.21 and so on. The margin error can be viewed as a measure of the variability of a large number of such simulated values; specifically, it is two times the standard deviation of a long sequence of simulated value. In the example of Charles's yards per carry, it is 1.16 yards.

We can interpret the margin of error in one of two ways. The most straightforward interpretation is based on the idea of repeating the "experiment" (e.g., a season, a game, or a career) and calculating a range of values that we expect to include the statistic of interest. For instance, in Charles's rushing example, if we were to repeat 2010 NFL season, we would expect Charles's yards-per-carry value to fall in the range

Real Estimate → Range of Error

$$6.38 \pm \underbrace{1.16}_{\text{Range of Error}} = (5.22,\ 7.54)$$

Therefore, the interval (5.22, 7.54) summarizes the variability in Charles's yards-per-carry value. A second interpretation, based on the underlying "true" characteristic of a player or team, is a little more subtle, but it is often more useful. Recall that a statistical estimate is our "best guess" of the characteristic of a probability distribution. However, we know that the estimate is not equal to that characteristic. For example, in the example for Charles, his observed average yards per carry is 6.38 yards; as noted, we expect this to be close to the true mean value of Y, but it is unlikely that it will be exactly equal to it. The observed value of 6.38 is an estimate of this true mean value.

The margin of error can also be interpreted in terms of how close we expect this hypothetical true mean value to be to the observed value of 6.38. The interval

$\bar{x}$ ne Confidence Interval (CI)

$$6.38 \pm 1.16 = (5.22,\ 7.54)$$

gives a range of values such that we are "reasonable certain" that Charles's true average yards per carry lies in that range. At this point, it is reasonable to ask why we are interested in the true mean yards per carry or in a range of results that might occur of the season is repeated. In fact, for some purposes we are not particularly concerned with these hypothetical values. This might be the case, for example, if we are just trying to summarize a player's or team's performance.

However, in other cases, we are primarily interested in understanding the process that generated the data to compare players or teams or to better understand what might happen in the future. In these cases, the underlying true values that are of interest and the variability of estimates are an important consideration both in understanding how future results might relate to the available data and in determining how much confidence we should place in the conclusions of the analysis. Either interpretation of the margin error gives us important information about the variability of estimates and their relationship to the hypothetical true value.

For example, suppose we are analyzing Kevin Durant's scoring for the 2011-2012 NBA season. Let X denote Durant's points scored for a given game, and suppose we model X as a random variable with some probability distribution. Although different aspects of this distribution might be of interest, depending on the context, here we focus on the mean of X, which we denote by μ. The available data are the points scored for the 66 games Durant played, and using the analogy principle, our estimate of μ is simply the average of the 66 values, 28, Durant's points per game (PPG) for the 2011-2012 regular season.

Although 28 exactly represents what actually occurred in the 2011-2012 season, it is only an estimate of μ. Durant's "true PPG," that is, the PPF he would achieve in a hypothetical long sequence of games. The uncertainty in this estimate is described by the margin error. Here the margin of error is 1.7; how this was determined is discussed in the following section. Therefore, our estimate of μ, Durant's true PPG for 2011-2012 season, could be reported as

$$28 \pm 1.7$$ (26.3, 29.7)

or as the interval (26.3, 29.7). That is, if we were somehow able to observe a long sequence of Durant's games, under the conditions of the 2011-2012 season, our best guess for his average PPG in this sequence is 28, and we are reasonably certain that it would fall between 26.3 and 29.7. Alternatively, we can use the interpretation of the margin of error in terms of repeating the season. That is, if we were able to repeat the 2011-2012 NBA season, we would expect Durant's PPG value to lie in the range 26.3 to 29.7.

There are two basic approaches to calculating the margin of error. The first is to use statistical formulas that are designed for this purpose. This approach is discussed in Section 4.3. The second approach is based on the interpretation of the margin of error in terms of hypothetical repetitions of the experiment. That is, we can compute the margin of error using compute simulation to obtain these hypothetical repetitions; this approach is discussed in Section 4.4.

4.3 Calculating the Margin of Error of Averages and Related Statistics

Although we will not generally calculate the margin of error by hand, using statistical formulas, it is useful to consider the margin-of-error formula in a few simple cases to better understand the issues that drive the accuracy of an estimate.

Consider the Durant scoring example from the previous example. Let X denote Durant's points scored for a given game, and suppose we are interested in μ, the mean of X. Our estimate of μ is the sample mean of Durant's point scored for the 66 games he played. The formula for the margin error in this case, in which we are estimating the mean of a random variable using a sample mean, is

Sample Standard Deviation

$$\frac{2S}{\sqrt{n}} \tag{4.1}$$

Sample Size

where S denotes the sample standard deviation of his game-by-game points scored and n denotes the number of data values used in the sample average. For example, $n = 66$ and S can be calculated by examining his game-by-game statistics, which show that $S = 6.9$. Thus, the margin of error is

ME: measuring the uncertainty of the sample average.

THE SMALLER IS THE ME
THE MORE ACCURATE THE MEASUREMENT.

ME =
Margin Error

$$\frac{2S}{\sqrt{n}} = \frac{2(6.9)}{\sqrt{66}} = 1.7$$

as reported in the previous section. The formula $\frac{2S}{\sqrt{n}}$ for the margin of error in this case shows the accuracy of the sample average as an estimate of the mean of a random variable is driven by two factors. One is the amount of information available for estimation, as reflected in the sample size n. The larger the sample size, the smaller the margin of error is. However, note that doubling the sample size does not cut the margin in half; because it is $\sqrt{n}$ ins the formula, not n, doubling the sample size corresponds to dividing the margin of error by about 1.4.

The other factor affecting the margin of error is the natural variability in the data, as measured by the standard deviation S. Therefore, measurements that are approximately constant have a small margin of error for estimating the mean, while measurements that show a lot of variability have a large margin of error. Table 4.1 shows the margins of errors for several statistics based on Durant's 2011-2012 season.

Table 4.1: Results from Durant's 2011-2012 season

Statistic	Mean	SD	Margin of Error
Rebounds	7.98	3.03	0.74
Assists	3.50	1.95	0.48
Turnovers	3.76	1.70	0.42
Fouls	2.02	1.36	0.34
Points	28.03	6.88	1.69

Although the formula for the margin of error of an average is based on the assumptions that we have a sequence of observations with the same mean and standard deviation, the result is still useful even if the means and standard deviations of the random variables vary, provided that we are interested in the overall average of the observations. In that case, $2S/\sqrt{n}$ tends to slightly overestimate the variability in the sample mean, but this overestimation is generally small.

For instance, in the Durant example, it may be appropriate to assume that the distribution of his points scored depends on whether the game is played at home or away, and in fact, his average PPG value is higher for home games, 28.8 PPG, compared to 27.2 for away games. Under this assumption, the margin of error for his overall scoring average (based on more sophisticated methods beyond the scope of this class) is 1.695; the margin of error using $2S/\sqrt{n}$ is 1.693. Hence, in this case, the margin of error based on $2S/\sqrt{n}$ is slightly less than the correct margin of error; however, the difference is negligible.

Another commonly used statistic with a simple formula for the margin of error is a proportion. Consider an experiment and let A denote the event of interest; let $\pi = P(A)$ be the probability of A occurring. Suppose that the experiment is performed n times and let X denote the number of times that A occurs. We have seen that X has a binomial distribution with parameters n and π. We can use X to estimate π. Define $p = X/n$; then, analogue of π, and it can be used as an estimator of π. The formula for the margin of error of this estimator is (PROBABILITY)

margin error formula

p: Estimated proportion
n: sample size.

$$2\sqrt{\frac{p(1-p)}{n}} \tag{4.2}$$

Note that this expression (4.2) has the same general form as the one for the margin of error of an average (4.1). The denominator is the amount of information as reflected in the sample size, and the numerator is a measure of the variability of the data. Because the data are particularly simple; for each experiment, either A occurs or it does not occur; the measure of variability is a function of the parameter estimate being analyzed.

Consider the following example: In 2012, Buster Posey had an on-base average (OBA) of 0.408 p based on 610 plate appearances. The margin of error corresponding to his OBA is therefore

Sample size: n

Single proportion

$$ME = 2\sqrt{\frac{0.408(1-0.408)}{610}} = 0.0398$$

so that we are reasonably certain that Posey's hypothetical true OBA for 2012 is in the interval $\boxed{(0.368, 0.448)}$. Note that, even with a sample size if 610 plate appearances, the margin of error is relatively large. This is often the case when estimating proportions. Table 4.2 gives the margin of error, based on an observed proportion of 0.40, for a wide range of sample sizes. Many statistics, such as batting average and OBA are traditionally reported to three decimal places. The results in Table 4.2 show that, even for a large sample size of 10,000 (roughly 15 seasons of at bats), the third digit in such a statistic is not particularly meaningful; this conclusion follows from the fact that the margin of error in that case is 0.005.

0.408 ± 0.0398

Table 4.2: Margins of error based on a proportion of 0.40

Sample Size	Margin of Error
500	0.022
1,000	0.015
2,000	0.011
5,000	0.007
10,000	0.005

There's always an ucertainty in an estimation

As with the case of the margin of error of an average, the formula for the margin of error of a proportion can also be used if the underlying probability of the event varies from trial to trial, provided that we are interested in estimating the overall average probability. For instance, in the Posey example, we might think that his on-base percentage (OBP) depends on whether the pitcher is right or left handed. This is supported by his 2012 results; his OBP against right-handed pitchers is 0.382, compared to 0.470 against left-handed pitchers. using a formula for the margin of error of his overall OBP that takes into account this difference (using methods beyond the scope of this course), the margin of error is 0.0396. Recall that the margin of error we calculated previously is 0.398, slightly higher but still very close to the correct value.

This discussion of the margin of error of a proportion applies when there are two outcomes of interest, such as "on base" and "not on base" for a given plate appearance. A similar formula applies when there are several possible outcomes. Suppose that for a given observation one of m events $A_1, A_2, \ldots, A_m$ occurs, and suppose that we are interested in a statistic that assigns weight w_j to event A_j. For instance, consider the slugging average of a MLB player, in which the events of interest are "single," "double," "triple," "home run," and "no hit." Slugging average assigns weight 1 to single, 2 to double, 3 to triple, 4 to home run, and 0 to no hit.

For a given set of data, based on n observations, let p_j denote the proportion of experiments in which A_j occurs. Then, the average value of the statistic for the dataset is given by

$$w_1p_1 + w_2p_2 + \cdots + w_mp_m$$

For example, consider Mike Trout in 2012. In his 559 at bats, he had 117 singles, 27 doubles, 8 triples, and 30 home runs. Therefore, if A_1 represents a single, $p_1 = 117/559 = 0.2093$; similarly, letting A_2 denote a double, $p_2 = 0.0483$; letting A_3 denote a triple, $p_3 = 0.0143$; and letting A_4 denote a home run, $p_4 = 0.0537$. It follows that his slugging average is

$$\overbrace{1(0.2093)}^{\text{Single}} + \overbrace{2(0.0483)}^{\text{Double}} + \overbrace{3(0.0143)}^{\text{Triple}} + \overbrace{4(0.0537)}^{\text{Home run}} = 0.564$$

Note that because no hits receives weight 0 in the slugging percentage calculation, we did not need to include p_5, the proportion of at bats in which Trout did not get a hit, in the calculation. For a statistic fo the form $T = w_1p_1+w_2p_2+\cdots+w_mp_m$, the formula for the margin of error is

Margin error Formula

$$2\sqrt{\frac{(w_1^2p_1 + w_2^2p_2 + \cdots + w_m^2p_m) - T^2}{n}} \tag{4.3}$$

multiple proportion

For example, for Trout, $T = 0.564$, then

The smaller the marging error is the more certain the data is.

$$w_1^2p_1 + w_2^2p_2 + \cdots + w_m^2p_m = 1^2(0.2093) + 2^2(0.0483) + 3^2(0.0143) + 4^2(0.0537) = 1.3904$$

and $n = 559$, so that the margin of error is

$$2\sqrt{\frac{1.3904 - 0.564^2}{559}} = 0.088$$

It follows that Trout's slugging percentage can be described as 0.564 ± 0.088 or as the interval $(0.476, 0.652)$. Note the margin of error for Trout's slugging percentage is much larger than the margin error for a batting average of 0.564 based on 559 at bats, which would be

$$2\sqrt{\frac{0.564(1-0.564)}{559}} = 0.042$$

This is because of the weighting that is used when calculating slugging percentage: By weighting home runs by 4, triples by 3, and doubles by 2, the natural randomness in the events home run, triple, and double is magnified. On the other hand, for batting average, every event receives either a weight of 1 (hits) or a weight of 0 (no-hits). This is generally true; statistics that are calculated by some event relatively large weight tend to have large margin of error. That is, there is often more uncertainty associated with the statistic than we might normally expect.

4.4 Using Simulation to Measure the Variation in More Complicated Statistics

In Section 4.2, an interpretation of the margin error was given in terms of hypothetical repetitions of the "experiment," such as a game or season, under consideration. The margin of error reflects the fact that if data are sampled from a given probability distribution by, for example, observing a player or team over the course of a season, the actual results observed will reflect the natural randomness of the process generating the data. In this section, we consider calculation of the margin of error by simulating this process. This approach is useful of more complicated statistics.

We begin by considering the example of Durant's per game scoring in 2011-2012. Although this statistic can be handled using the techniques of the previous section, it is instructive first to apply the simulation-based method to a simple case before considering more complicated ones.

Let X denote his points scored in a given game and, based on data from the 2011-2012 season, the average value of X is 28.0. The margin of error of this estimate measures the natural variability in this estimate. That is, if we were somehow able to replay the 2011-2012 NBA season, Durant's game-by-game points scored values, as well as his season average PPG would be different, reflecting this natural variability.

To understand the variability in his average PPG, we can "simulate" a 2011-2012 NBA season, or at least Durant's scoring for that season. To do this, we can simulate 66 points-scored values from a probability distribution that represents Durant's scoring. Based on these values, we can compute an average PPG value for this simulated season. If we repeat this process many times, we obtain simulated average PPG values

$$\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_M$$

where M denotes the number of simulated seasons. The variation in these simulated averages gives us some information about the variability of Durant's average PPG; specifically, the margin of error is given by $2\hat{S}$, where $\hat{S}$ denotes the sample standard deviation of $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_M$.

The 66 PPG values from Durant's 2011-2012 season are given in Table 4.3. To sample form this distribution, we simply randomly choose 66 new values from the values in Table 4.3. To do this, we could choose a random integer from 1 to 66 and choose the corresponding value from Table 4.3. For instance, if the random number is 6, Durant's points scored for the first game in his simulated season is 24; if the second random number is drawn is 56, Durant's points scored for the second game in his simulated season is 22, and so on.

Simulated dataset

Table 4.3: Durant's game-by-game scoring in 2011-2012

32	32	35	29	29	24	43	29	22	19	23	44	30	21	26	21	25	28
40	32	18	26	25	24	28	26	23	30	22	35	38	23	33	28	31	51
23	33	21	19	27	33	33	22	36	23	36	37	25	20	20	33	28	28
29	22	21	27	26	19	27	12	30	32	33	30						

Suppose we repeat this process until we have 66 values, corresponding to an entire simulated 2011-2012 season. An example of such a dataset is given in Table 4.4. The average of the 66 values in Table 4.4 is 27.7. As expected, this is close to Durant's actual scoring average for 2011-2012, but it is not exactly the same; the difference between the values gives us some indication of the variability in scoring averages over a 66 game season.

Table 4.4: Simulated game-by-game scoring by Durant

24	22	27	24	23	26	22	38	28	40	25	28	37	35	35	29	22	24
19	20	22	28	19	25	26	32	25	21	22	30	28	26	36	26	33	21
33	35	26	32	29	21	35	21	28	19	19	40	33	28	26	27	40	38
27	22	32	22	33	29	19	38	25	29	22	33						

More information about this variability can be obtained by simulating several 2011-2012 seasons for Durant. As discussed previously, let $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_M$ denote the simulated scoring average for M simulated seasons; therefore, $\overline{X}_1 = 27.7$. I repeat the entire procedure nine times so that we have the results of 10 simulated seasons; Durant's simulated scoring averages for these seasons are 27.7, 28.8, 27.9, 26.9, 27.0, 27.1, 29.6, 28.7, 28.9, 28.5. The mean of these values is 28.1, which is close to Durant's actual PPG for the 2011-2012 season. However, it is the variability of these values that is of main interest to us.

These values show how we might expect the average PPG of a player with Durant's skill level, for a 66-game season, to vary. The values are all relatively close to Durant's actual average PPG of 28.0, but they range from 26.9 to 29.6. The standard deviation of these values is 0.92, and twice this standard deviation, 1.84, can be used as the margin of error for his average PPG. This is close to, but not exactly equal to, the value we obtained using the margin of error formula (1.69). The advantage of the simulation-based approach is that we do not need to know the formula to calculate the margin of error.

Of course, it is important to recognize that the value 1.84 is based in the random numbers used in the simulation. I repeat the entire process, simulating 10 more seasons and computing the margin of error based on the results; this procedure yielded the margin of error 1.55, again close to the formula-based result but not exactly equal to it. The following box shows the R code that was used to simulated Durant's game-by-game scoring in 2011-2012.

R code

```
## Durant's game-by-game scoring 2011-2012 season
durant = c(32, 32, 35, 29, 29, 24, 43, 29, 22, 19, 23, 44, 30, 21, 26, 21, 25, 28,
           40, 32, 18, 26, 25, 24, 28, 26, 23, 30, 22, 35, 38, 23, 33, 28, 31, 51,
           23, 33, 21, 19, 27, 33, 33, 22, 36, 23, 36, 37, 25, 20, 20, 33, 28, 28,
           29, 22, 21, 27, 26, 19, 27, 12, 30, 32, 33, 30)

## Simulating game-by-game scoring
results = array()

for(i in 1:1000){

   ## Here we simulate the data
   sim_data = durant[sample(length(durant), replace = T)]

   ## Computing and storing the average
   results[i] = mean(sim_data)
}

## Computing the average
mean(results)

## Computing the standard deviation of the averages
sd(results)
```

The following box shows the Python code that was used to simulated Durant's game-by-game scoring in 2011-2012.

Python code

```
import numpy as np
from sklearn.utils import resample

## Durant's game-by-game scoring 2011-2012 season
durant = np.array([32, 32, 35, 29, 29, 24, 43, 29, 22, 19, 23, 44, 30, 21, 26,
                   21, 25, 28, 40, 32, 18, 26, 25, 24, 28, 26, 23, 30, 22, 35,
```

Python code

```
                    38, 23, 33, 28, 31, 51, 23, 33, 21, 19, 27, 33, 33, 22, 36,
                    23, 36, 37, 25, 20, 20, 33, 28, 28, 29, 22, 21, 27, 26, 19,
                    27, 12, 30, 32, 33, 30])

## Simulating game-by-game scoring
results = list()

for i in range(0, 1000):

    ## Here we simulate the data
    sim_data = resample(durant, replace = True)

    ## Computing and storing the average
    results.append(np.mean(sim_data))

## Computing the average
np.mean(results)

## Computing the standard deviation of the averages
np.std(results, ddof = 1)
```

4.5 Comparison of Teams and Players

Analytic methods are often concerned with comparing and contrasting players or teams. In making such comparisons, it is important to take into account the inherent variability in sports data. For example, suppose we are comparing the scoring of Durant and LeBron James in 2011-2012 NBA regular season. According to game results from that season, Durant average 28.0 PPG, while James averaged 27.1 PPG. Based on these results, Durant clearly had a per game scoring average 0.9 PPG higher than that of James. However, can we conclude that Durant is a better scorer than James? Or, could the fact that Durant had a higher scoring average than James be because of the random nature of scoring in the NBA, and if we observed a very long sequence of games for both players, played under the conditions of the 2011-2012 season, would James actually have the higher scoring scoring average?

To address this issue, we can take into consideration the margin of error of the 0.9 PPG difference in Durant's and James' scoring averages. Using the same approach used for Durant, the margin of error for James' scoring average is also 1.7. To find the margin of error for the difference between Durant's and James' scoring average, we use the following rule: The margin of error for the difference between two independent measurements is found by adding the squares of the two margins of error and taking the square root of the result. It follows that for the difference between Durant's and James' scoring average, the margin error is

$$\sqrt{1.7^2 + 1.7^2} = 2.4$$

and the difference between Durant's and James' scoring average can be reported as 0.9 $\pm$ 2.4 PPG. That is, our best guess of the true difference is 0.9 PPG, and we are reasonably certain that the true difference lies between -1.5 and 3.3 PPG. Note that this range includes negative values, suggesting that the true difference in scoring might favor James. Table 4.5 shows comparisons of other statistics of Durant and James in the 2011-2012 NBA season.

Table 4.5: Comparison of Durant and James based on 2011-2012 per game averages

	Durant	James	Difference	Margin of Error
Rebounds	8.0	7.9	0.1	1.0
Assists	3.5	6.2	-2.7	0.8
Turnovers	3.8	3.4	0.3	0.6
Fouls	2.0	1.5	0.5	0.5
Points	28.0	27.1	0.9	2.4

When comparing Durant's and James' statistics, we have chosen to base the comparisons on differences of their per game averages. An alternative approach is to consider the ratio of one player's average to the other player's average. For instance, to compare per game scoring, we could compute the ratio of Durant's PPG to James' PPG:

$$\frac{28.0}{27.1} = 1.033$$

so that Durant's PPG is 3.3% higher than James'. Note that such an approach is generally only used for non-negative statistics in which the average values are positive. As with differences, it is important to take into account the margin of error of the ratio. Unfortunately, the formula for the margin of error of a ratio is a little more complicated than the formula for the margin of error of a difference. Consider two measurements X and Y; let $\overline{X}$ and $\overline{Y}$ denote the corresponding sample means and let $R = \overline{X}/\overline{Y}$ denote the ratio. Then, the margin error of R is given by

$$R\sqrt{\left(\frac{ME(\overline{X})}{\overline{X}}\right)^2 + \left(\frac{ME(\overline{Y})}{\overline{Y}}\right)^2} \tag{4.4}$$

where $ME(\overline{X})$ and $ME(\overline{Y})$ denote the margins of error of $\overline{X}$ and $\overline{Y}$, respectively. Therefore, in the Durant and James example, the margin of error of the ratio is

$$(1.033)\sqrt{\left(\frac{1.7}{28.0}\right)^2 + \left(\frac{1.7}{27.1}\right)^2} = 0.090$$

That is, the ratio of Durant's PPG to James' PPG is 1.033 with a margin of error of 0.090 so that the hypothetical true ratio is in the range 0.943 to 1.123; therefore, as with the difference-based analysis, the scoring averages of these two players are essentially the same.

When comparing parameter values, the comparison can be based either on differences or ratios, depending on which approach seems more natural for the quantity analyzed. Because scoring averages are more commonly compared by noting how many more PPG one player scores than another, in the Durant-James scoring example, a comparison based on differences seems more appropriate.

It is worth noting that when comparing parameters through their differences, it does not matter how the difference is calculated. For instance, Durant's PPG minus James' PPG is 0.9 with a margin of error of 1.4. If we look at James' PPG minus Durant's PPG, the result is -0.9 with a margin of error of 1.4, so we reach the same conclusion in either case. However, with ratios, the situation is a little different. The ratio of James' PPG to Durant's PPG is 0.968 with a margin of error of 0.084, so the true ratio is in the range 0.884 to 1.052. If we invert these values to obtain a range for the Durant-to-James ratio, the result is 0.951 to 1.131, which is close to, but not exactly the same as, the result obtained when we analyzed the Durant-to-James ratio directly. Fortunately, in many cases, the differences between the two analyses are small and, as in this case; do not affect the general conclusions.