MMA Analysis

Ian Young

June 2023

Contents

Data Introduction

The data used for the following examples is MMA fight data. The data has 128,946 observations and 30 variables. The data includes fights from 1991 to 2023. An event may have multiple fights within it, each fight has two observations for each fighter participating in the fight. Along with the general information of the fight such as date and event, there is information on the fight result and fighter information. Various methods may result in the data being filtered down, and the removal of NA's.

Listed below is a data dictionary.

Variable	Definition	Example_Values
stack	number represents fighter 1 or fighter 2 in original dataset	1 or 2
date	date of the event (mdy)	10/12/2023
event_day	day of the event	12
event_month	month of the event	10
event_year	year of the event	2023
event	name of the event	UFC 1: The Beginning
champ	is the event a championship event	TRUE/FALSE
decision	fight result	TKO (Injury)
decision_group	aggregated groups of decision variable	KO/TKO
round	how many rounds the fight lasted	3
time	time the fight ended	4:42
time_minutes	time the fight ended in minutes	4.7
time_seconds	time the fight ended in seconds	282
fighter_result	result of the fight for the fighter	W
fighter_id	a unique ID for each fighter	2354059
fighter	full name of the fighter	Shamrock, Ken
first_name	first name of the fighter	Ken
last_name	last name of the fighter	Shamrock
Nickname	nickname of the fighter	The World's Most Dangerous Man
fighter_country	home country of the fighter	USA
sex	sex of the fighter	M
Height	height of the fighter in inches	61
Weight	weight of the fighter in pounds	155
fighter_age	age of the fighter at the time of the fight	23
Reach	reach (arm span) of a fighter measured in inches	68
Stance	stance the fighter uses	Orthodox
WT Class	specific weight class the fighter was in at the event	Heavyweight - Old UFC
weight_class	aggregated groups of WT Class	Heavyweight
Team	team the fighter was a part of at the time of the fight	Long Island MMA
Birthdate	birthday of the fighter	12/22/1989

Chi-square Test Overview

The chi-square test is a test used to determine if there is a significant association between two categorical variables. It compares the observed frequencies of the different categories with the expected independent frequencies. It can also be thought of as determining if dimensions of a contingency table are associated. Lets use the MMA data to determine if since the year 2020 fighter weight class and the decision of a fight are associated. After selecting, filtering, and preparing the data, a contingency table is produced (Note: To perform a chi-square test in many software packages, it is often not necessary to format the data into a contingency table. Many packages can take data from a standard rectangular data table and pass it through a chi-square function.). Additionally, it may be of use to present the data in a more visual format.

Category	Heavyweight	Middleweight	Sum
Decision	222	173	395
KO/TKO/Submission	559	346	905
Sum	781	519	1300

Chi-square 2x2 Example

The question we are trying to answer is, is there an association between weight class and fight decision?

This leaves us with the following null and alternative hypothesis...

H_0: Weight class and the fight decision are not associated.
H_a: Weight class and the fight decision are associated.

Category	Heavyweight	Middleweight	Sum
Decision	222	173	395
KO/TKO/Submission	559	346	905
Sum	781	519	1300

Decision: 222 of 395
KO/TKO/Submission: 559 of 905

If there was no association we would expect the same proportion of Decision & KO/TKO/Submission in the heavyweight class.

Our best guess to this common proportion of heavyweight is (559 + 222) / (395 + 905) = 0.601

Now using this common proportion we estimate our best guess at the table counts if there were no association (also known as the expected counts).

Expected heavyweight = number in group x 0.601
Expected middleweight = number in group x (1 - 0.601)

Another way to find the expected count for each cell is the following formula:
expected count = (row total x column total) / grand total

For example we can compute the expected count for the first cell as follows...
(395 x 781) / 1300 = 237.3

The expected counts are then used to construct the test statistic. Below is a table of the expected counts.

	Heavyweight	Middleweight
Decision	237.3	157.7
KO/TKO/Submission	543.7	361.3

The test statistic takes the difference of the observed minus the expected and squares this number and divides the number by the expected count. This is done for each cell and all the numbers are summed together to get the test statistic. Below is an example using our data.

```
(((222-237.3)^2) / 237.3) +
  (((559-543.7)^2) / 543.7) +
  (((173-157.7)^2) / 157.7) +
  (((346-361.3)^2) / 361.3)
```

```
## [1] 3.549334
```

We Can also calculate the test statistic, degrees of freedom, and p-value using software packages. We will use an alpha of 0.05 as our cutoff.

statistic	chisq_df	p_value
3.322884	1	0.0683219

The test returned a p-value of 0.0683. Our conclusion then after performing a chi-square 2x2 test is at $p < 0.05$ we do not have evidence to suggest that weight class and fight decision are associated ($p = 0.0683$). In other words, we fail to reject our null hypothesis.

Fisher's Exact Test Example

The chi-square test only works if expected counts/frequencies are large enough. A general guideline is that chi-square test can be used if most of the expected counts are at least 5. If this is not the case, the "Fisher's Exact Test" is an alternative test. The Fisher's Exact Test assumes the row and column totals are fixed, then calculates the probability of a particular set of cell counts, then figures out all possible tables that could be constructed given the fixed totals, and calculates probabilities for each of these tables. The test determines which of these tables are extreme or more extreme than the table seen, and sums up the probabilities for the table seen and the more extreme tables to get the p-value.

For example, lets say we are looking in the data at fights from 2023. We want to look and see if there is an association between the two lightest weight classes (flyweight and strawweight) and the decision type. We first check the expected counts in each cell.

	Decision	KO/TKO/Submission
Flyweight	4.7	10.3
Strawweight	1.3	2.7

We notice that 3 of the 4 cells have expected counts less than 5, so we will use Fisher's Exact Test in place of the Chi-Square test.

We once again state our hypotheses.

H_0: Weight class and the fight decision are not associated.
H_a: Weight class and the fight decision are associated.

```
##
##  Fisher's Exact Test for Count Data
##
## data:  dat2$p1_wtClass2 and dat2$decision_group
## p-value = 0.5573
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.02083739 7.00380583
## sample estimates:
## odds ratio
##  0.3859077
```

The test returned a p-value of 0.5573. Our conclusion then after performing a Fisher's Exact Test is at $p < 0.05$ we do not have evidence to suggest that weight class and fight decision are associated ($p = 0.5573$). In other words, we fail to reject our null hypothesis.

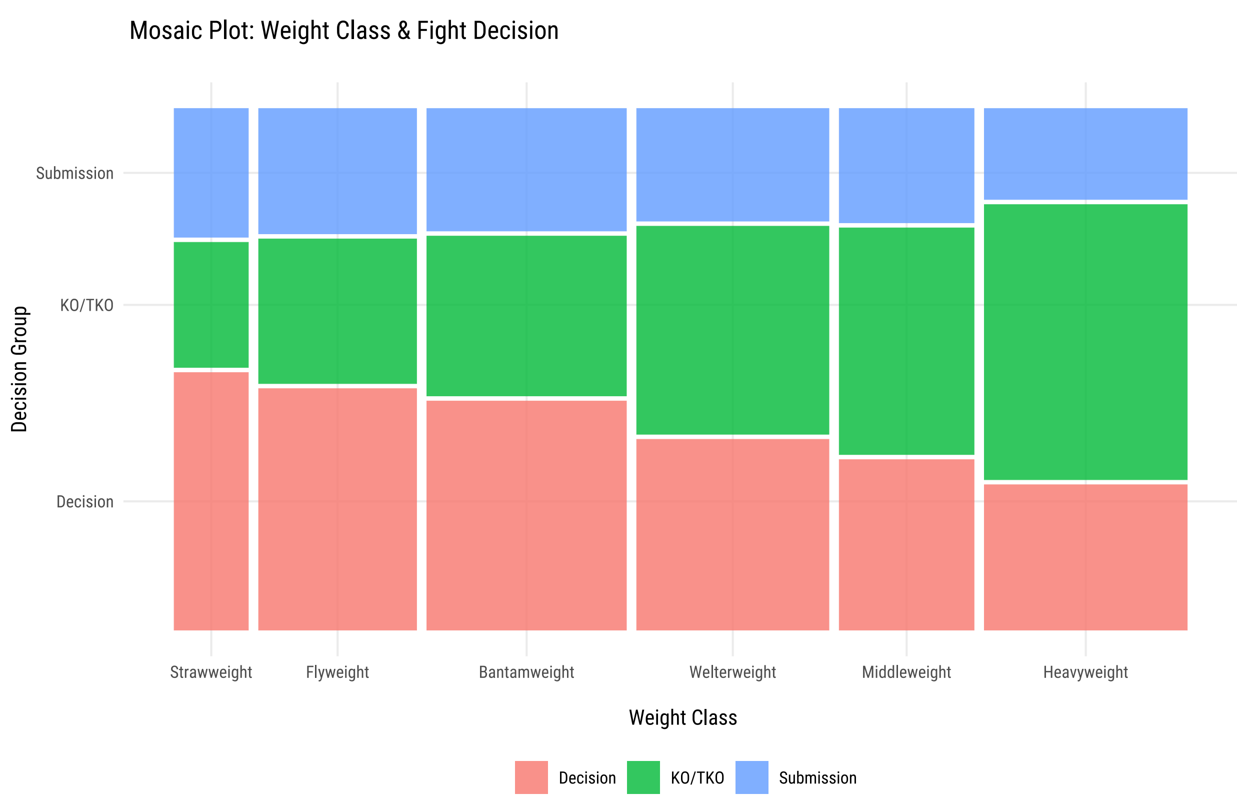
Chi-Square rxc Example

The Chi-Square 2x2 example we did earlier, can be generalized on a table of any number of rows (r) and columns (C). Lets take a look at a Chi-Square Test on a larger table. Below we see a table of 6 columns representing weight classes and 3 rows representing categories. We have a 3 x 6 table of observed values followed by the corresponding table of expected values.

Category	Strawweight	Flyweight	Bantamweight	Welterweight	Middleweight	Heavyweight	Sum
Decision	144	287	344	276	173	222	2236
KO/TKO	70	172	240	301	230	420	2040
Submission	72	149	184	163	116	139	1230
Sum	286	608	768	740	519	781	5506

	Strawweight	Flyweight	Bantamweight	Welterweight	Middleweight	Heavyweight
Decision	111.7	237.5	300.0	289.0	202.7	305.1
KO/TKO	110.7	235.3	297.3	286.4	200.9	302.3
Submission	63.6	135.2	170.7	164.5	115.4	173.6

As we saw above with the 2x2 Chi-Square test it may be helpful to display this information visually. Below we plot what is called a mosaic plot. On the x axis is weight class and the wider the bar the higher the count in that weight class and the thinner the bar the lower the count in that weight class. On the y axis we see the different decision groups, again the width shows how many in that particular decision group. Looking at the overall plot we can see how the proportion of weight class and decision groups vary.



Lets formally state our null and alternative hypothesis. Again the 2x2 Chi-Square test generalizes to the rxc Chi-Square test so the following will feel very similar.

H_0: Weight classes and the fight decisions are not associated.
H_a: Weight classes and the fight decisions are associated.

We then proceed to compute our test statistic, degrees of freedom, and p-value. We will use an alpha of 0.05 as our cutoff.

statistic	chisq_df	p_value
157.9847	10	8.442328e-29

The test returned a statistic of 157.9847 with a resulting p-value of <.0001 and 10 degrees of freedom. Our conclusion then after performing a chi-square 3x6 test is at $p < 0.05$ we do have evidence to suggest that weight classes and fight decisions are associated ($p < .001$). In other words, we reject our null hypothesis.

However, this still begs a couple of questions. (1) Which weight classes and which decision groups differ? (2) We achieved statistical significance (maybe as a result of high counts), but is there practical significance?

Lets do an informal inspection of the observed and expected values and look into the residuals (difference between the observed and the expected counts).

Below are the observed values.

	Strawweight	Flyweight	Bantamweight	Welterweight	Middleweight	Heavyweight
Decision	144	287	344	276	173	222
KO/TKO	70	172	240	301	230	420
Submission	72	149	184	163	116	139

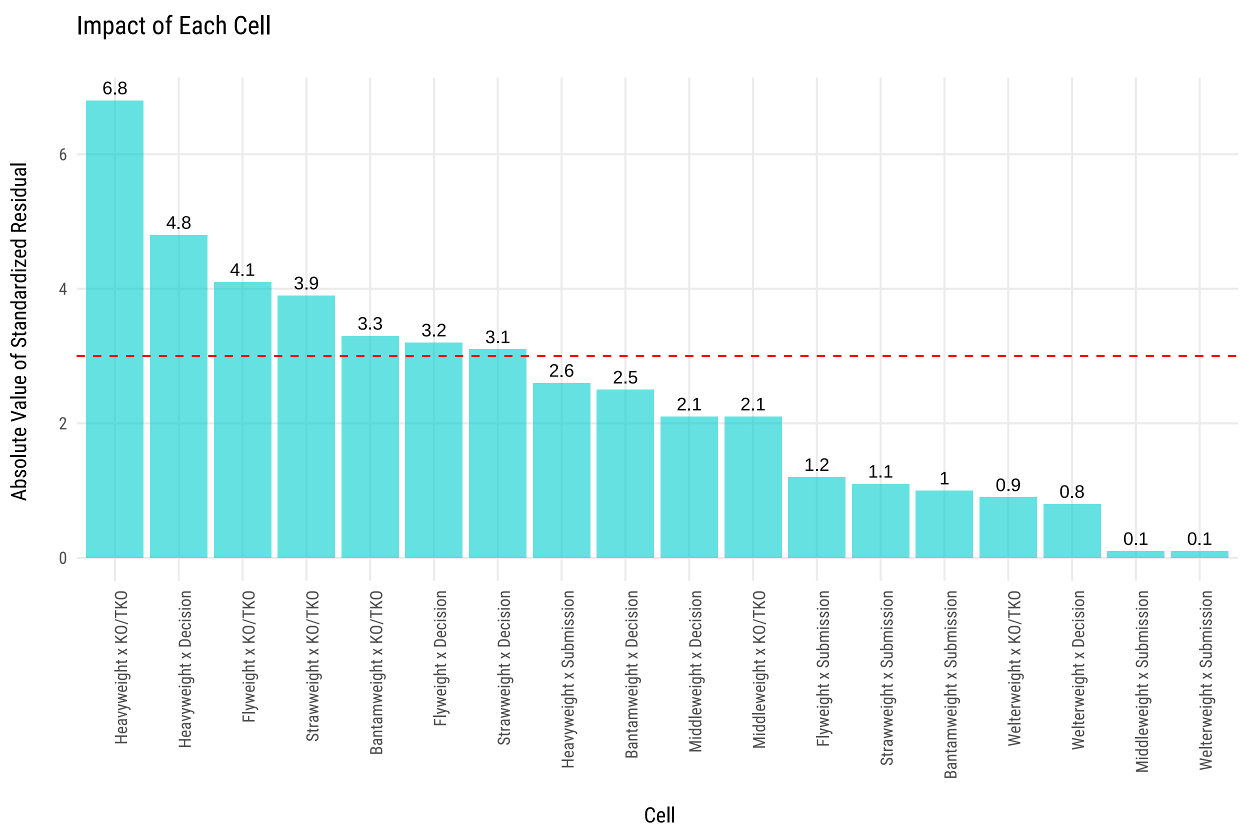
Below are the expected values.

	Strawweight	Flyweight	Bantamweight	Welterweight	Middleweight	Heavyweight
Decision	111.7	237.5	300.0	289.0	202.7	305.1
KO/TKO	110.7	235.3	297.3	286.4	200.9	302.3
Submission	63.6	135.2	170.7	164.5	115.4	173.6

With this high of rows and columns it is difficult to see if there are practical differences between certain cells. This may be achievable with a 2x2 table, but is much more difficult with larger tables. To help use we can calculate the standardized residuals to see cells to investigate. The formula for the Pearson residuals is (O-E)/sqrt(E).

	Strawweight	Flyweight	Bantamweight	Welterweight	Middleweight	Heavyweight
Decision	3.1	3.2	2.5	-0.8	-2.1	-4.8
KO/TKO	-3.9	-4.1	-3.3	0.9	2.1	6.8
Submission	1.1	1.2	1.0	-0.1	0.1	-2.6

Now we can look at the standardized residuals and see which cells may be impacting the test and causing a conclusion that there is an association. However, we still run into the problem of trying to look at many different cells when the table is large. To help us, let's plot the absolute values of the standardized residuals and sort them.



Looking at the plot we can now see certain cells that had high standardized residuals (above 3), with some very large residuals in the heavyweight group impacting the test.

Chi-Square rxc Exercise