

MLBにおけるトミー・ジョン手術発生要因の実証分析

t 検定・ロジットモデルを用いたリスク要因の定量化

慶應義塾大学 経済学部 4 年

長倉大輔研究会

宮崎 浩一

2015 年 12 月

要旨

本論ではメジャーリーグ・ベースボールにおけるトミー・ジョン手術件数の急激な増大たる状況に対して、実証分析による手術の発生要因の検証を行った。トミー・ジョン手術が発生する要因として本論では「ストレートの球速向上」「特定球種の使用が与えるヒジへの負担」「メカニクス上の欠陥」「身体的差異」の計 4 つの仮説を紹介しているが、そのいずれもが有力である一方で決定的な要因とは言えない。そこで本論では t 検定、ロジットモデルを使用することで、これら諸説の検証と同時に、トミー・ジョン手術の要因の影響度合を数値として可視化する試みを行った。その結果、本論では特に「イニング数」「ストレートの平均球速」「ストレートの使用割合」「チェンジアップの平均球速」の 4 点が重大な要因であるとの結論を得た。

目次

1. 本論目的
2. TJS 要因に関する仮説・先行研究の参照
 - 2.1 仮説
 - 2.1.1 MLB における球速向上
 - 2.1.2 特定球種の使用が与えるヒジへの負担
 - 2.1.2.1 スプリッター
 - 2.1.2.2 カッター
 - 2.1.3 メカニクス上の欠陥
 - 2.1.4 身体的差異
 - 2.2 先行論文
 - 2.2.1 Andrews et al (2006) における実証分析アプローチの参照
 - 2.2.1.1 t 検定・ χ^2 検定
 - 2.2.1.2 ロジットモデル
 - 2.2.2 Andrews et al (2006) の分析結果
 - 2.2.2.1 t 検定・ χ^2 検定
 - 2.2.2.2 ロジットモデル
 - 2.2.3 Andrews et al (2006) から本論が参照すべき点
3. 基本統計量と実証分析手法
 - 3.1 基本統計量
 - 3.2 分析手法
 - 3.2.1 t 検定
 - 3.2.1.1 等分散における t 検定
 - 3.2.1.2 異分散における t 検定 (Welch 検定)
 - 3.2.1.3 本論における t 検定の利用
 - 3.2.2 ロジットモデル
 - 3.3 本論における分析の手順
4. 分析結果

4.1 t 検定

4.2 ロジットモデル

4.2.1 単一変数によるロジットモデル

4.2.2 ステップワイズ法により採択されたロジットモデル

5. インプリケーション・課題

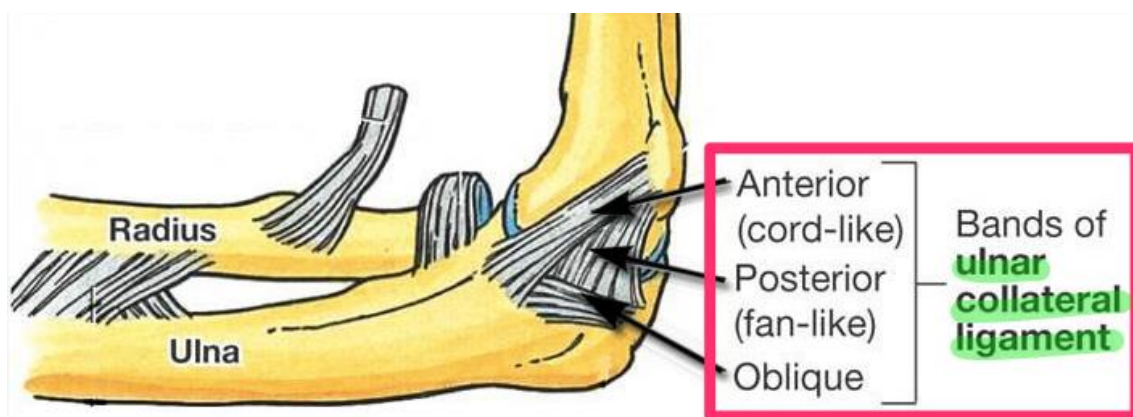
参考文献・データ

1. 本論目的

トミー・ジョン手術（以下、TJS）とは、医学用語上で「側副靱帯再建術」と言われる外科的処置の1種であり、特に野球たるスポーツにおいては主に投手の症例が見受けられる。

下図-1 内における“Ulnar Collateral Ligament”（内側側副靱帯 以下、UCL）が損傷・断裂を負った場合に、そのリカバリーに向けて大腿裏・臀部・ヒザ等の箇所から靱帯を採取し、損傷したものに代わる靱帯として移植を行う術式である。1974年にアメリカの整形外科医であるフランク・ジョーブ（1925～2014）が当時メジャーリーグ・ベースボール（以下、MLB）のロサンゼルス・ドジャースでプレーしていた投手 Tommy John に対し実施したのが史上初の症例であり、“@MLBPlayerAnalys”のリサーチ結果によると2015年8月現在までTJSを経験したプレーヤーはメジャー・マイナー問わずのべ1081人を上回る¹。

図-1 ヒジにおける内側側副靱帯



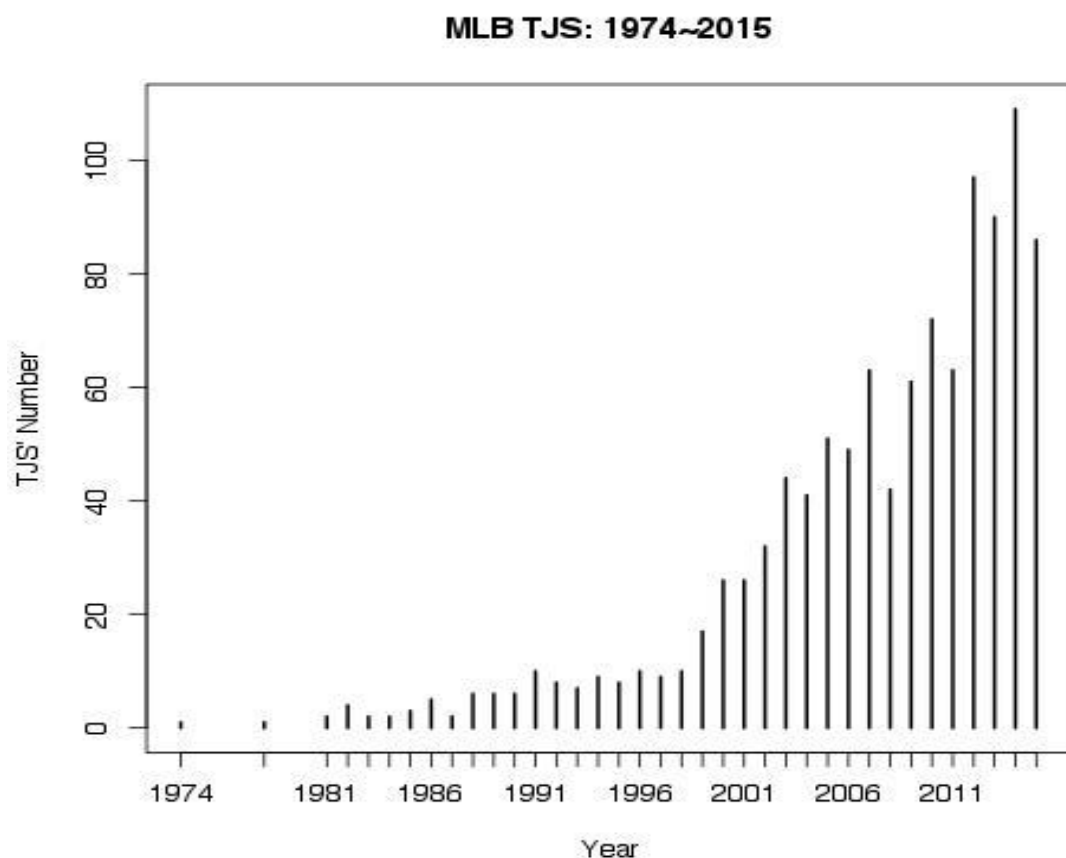
参照：<http://ishareimage.com/ulnar-collateral-ligament-elbow.asp>

“@MLBPlayerAnalys”曰く、TJSの症例はここ10～15年で大きな増加を辿っている。上記では1974～2015年3月までにのべ1081人のプレーヤーが実行したことを述べたが、うち752人が2006～2015年の10年間での症例である。次ページ図-2は、各年のTJS症例数の推移を注釈1のデータより作成したグラフである。このグラフからは、特に2000年代に入ってからTJS症例の大幅な増加を明らかに読み取ることが出来る。また、表-1に

¹ Tommy John Surgery List (@MLBPlayerAnalys), <https://docs.google.com/spreadsheets/d/1gQuiXQQGOVNaiuwSN680Hq-FDVscwvN-3AazykOBON0/edit#gid=0>, アクセス日：2015年8月10日

は年度による増加の傾向について OLS 推定を行った結果を示しているが、1974～1999 年と以降の期間では 1 年あたりでの症例増加割合に大きな差が存在することを捉えることが出来る。

図-2 年度別での TJS 症例の推移



データ参照：Tommy John Surgery List (@MLBPlayerAnalysis),
<https://docs.google.com/spreadsheets/d/1gQujXQQGOVNaiuwSN680Hq-FDVSCwvN-3AazykOBON0/edit#gid=0>, アクセス日：2015 年 8 月 10 日

表-1 TJS 症例数推移の OLS 推定結果

データ期間	Intercept	1 年での症例増加数推定
全体	-22.3693***	2.2364***
1974～1999	-2.1293**	0.5224***
2000～2015	17.400**	4.953***

注：“****”は0.1%有意、“***”は1%有意を示す

特に史上初めて症例数が90を上回った2012年から、MLBはTJSの要因究明に尽力している。各クラブチームでは故障を防ぐ上で有効な投手運用策を検討・実施しており、特に予め1年間でプレーするイニング数、もしくは投球数を定めた上で、その上限に達した時点でシーズンでのプレーを強制的に終了させる管理手法はMLBでも最もオーソドックスなアプローチとなっている。また、スポーツ医学のシンクタンクとして1986年に設立された「アメリカスポーツ医学研究所」(以下、ASMI)でも運動力学からのアプローチを用いてTJSの要因についての提言が行われており、各所でのTJS増加抑止に向けた取り組みが進んでいる。

本論ではTJSにおける上記の状況を受けて、グループ間での差異分析、及びロジットモデルを用いたTJSの要因分析を行った。TJSの発生要因についてはMLBでも今日に至るまでいくつかの仮説が提唱されているが、決定的となる要因を得るには至っていない。実証分析によるTJSリスクの可視化は上述した増大するTJS件数の抑制において重大な意味を有している。特に「TJS経験の有無による区分した2グループ間に存在する何らかの差異」「TJS発生確率の増大に作用するリスク要因の決定」は重要なトピックであり、本論の目的はこれら2点に対する統計学的見地からの一定の結論を得ることにある。

ただし、後述する仮説において、その根拠を統計学的な領域以外に求める視点が存在する点から、本論では実証分析が可能な部分においてのみ検討を行った。すなわち、実証分析においてカバー出来なかった点は別に言及が求められる部分であり、本論はその示唆を行うにとどまる内容である。

本論の構成は以下の通りである。第2章ではTJSの要因についての各仮説、及びASMIによる故障要因に関する実証分析の先行研究を参照した。第3章では本論にて使用する分析手法、及び基本統計量について整理した。第4章では第3章で示したモデルによる分析結果を提示し、TJS要因について言及を行った。第5章ではTJSの要因について妥当性の検討を実施した上で、本論の最終的な結論を導いた。

2. TJS 要因に関する仮説・先行研究の参照

本章では TJS 要因について MLB で検討がされている仮説、及び ASMI が 2006 年に行った故障要因に関する実証分析の先行研究を参照する。

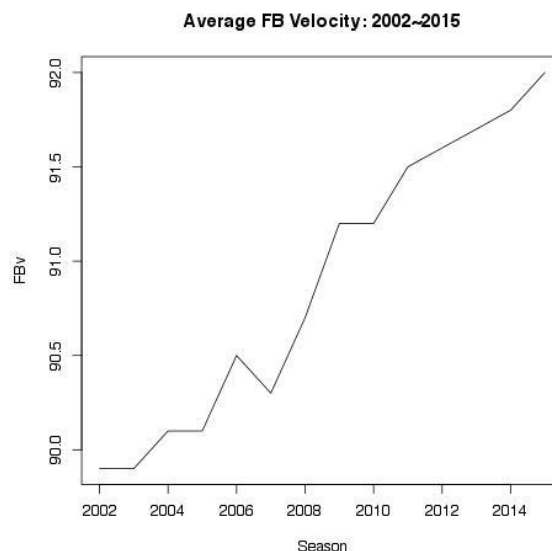
2.1 仮説

本節では MLB において TJS 要因として検討されている仮説を参照する。ここでは計 4 本の仮説を紹介するが、各説を支持する根拠は統計学に限らず、医学・動力学・経験則等幅広い領域に及ぶ点を確認されたい。

2.1.1 MLB における球速向上

2000 年に入ってから TJS 増大の要因に関する各仮説で最も大きな支持を集めている仮説として、MLB における球速向上がある。MLB データを取り扱う米シンクタンクサイトの “FanGraphs” (<http://fangraphs.com>) 曰く、MLB 全体の球速計測をスタートした 2002 年から今日に至るまで、リーグの平均球速は増大傾向にある。2015 年のストレートの平均球速は 92.0 マイルを示しており、これは計測史上トップの値である（下図-3 参照）。

図-3 2002～2015 年の MLB におけるストレートの平均球速の変化



参照：FanGraphs.com,

<http://www.fangraphs.com/leaders.aspx?pos=all&stats=pit&lg=all&qual=0&type=4&season=2015&month=0&season1=1974&ind=0&team=0,ss&rost=0&age=0&filter=&players=0>, アクセス日：2015 年 8 月 10 日

なお、図-3 のグラフの OLS 推定結果は下表-2 の通りである。この結果は、MLB においては 2002 年から 1 年を経る度におおよそ 0.17867 マイルの球速向上が認められたことを示している。

表-2 図-3 の OLS 推定結果

Intercept	Per Year (1 年での球速変動値)
89.731	0.17867

ASMI は 2014 年 7 月に “Position Statement for Tommy John Injuries in Baseball Pitchers” たる TJS 要因についてのレコメンデーションを HP にてアップし、そこでは以下のように言及している²。

“ Vary speeds for each of your pitch types. This will not only reduce the overuse on the elbow, but also can be an effective strategy. The best professional pitchers pitch with a range of ball velocity, good ball movement, good control, and consistent mechanics among their pitches. The professional pitcher’s objectives are to prevent baserunners and runs, not to light up the radar gun.

Pitchers with high ball velocity are at increased risk of injury. The higher the ball velocity, the more important to follow the guidelines above.”

ここでは特に若いプレーヤーにおける球速に対するこだわりの払拭を一貫して主張しており、速い球速が故障リスクの増大に結び付くことを危惧している。引用内にて ASMI は、MLB でプレーする投手として成功を収める上では、前提として球速を一定レベルに維持しながら「ボールのムーブメント」「コントロール」「メカニクスの維持」の 3 点を実現することが重要であるとの結論に至っている。この ASMI の見立ては特に医学的見地上におけるスタンダードであり、テキサス・レンジャーズのチームドクターである Keith Meister は「フォルクスワーゲンに V8 エンジンを搭載出来ないことと同じである。(ストレートの速

² ASMI, “Position Statement for Tommy John Injuries in Baseball Pitchers”, <http://www.asmi.org/research.php?page=research§ion=TJpositionstatement>, アクセス日：2015 年 8 月 11 日

さを不用意に追い求めることは) 最終的に身体を壊してしまう」³とのコメントで ASMI の主張を支持している。

2.1.2 特定球種の使用が与えるヒジへの負担

主にスプリッターとカッターの利用について、MLB では UCL に大きな負担を与える要因として捉えている。

2.1.2.1 スプリッター

スプリッターは図-4 で示されるように、人差し指・中指を広げるグリップから投げられる球種である。動力学上ではこのグリップが UCL への負担をより直接的にすると論じられており、MLB ミネソタ・ツインズのピッチングコーチである Rick Anderson も “You can just take your fingers and the more you put them apart, the more you put stress on the elbow” と、グリップを由来とした負担の増大について言及している⁴。

図-4 スプリッターのグリップ



参照：週刊ベースボールオンライン, http://column.sp.baseball.findfriends.jp/?pid=column_detail&id=001-20140428-04

一方で図-2 に示される TJS の増大要因として捉える上では、スプリッター自体が TJS の急速な増大が進行した 2000 年以降で決して主立って使用された球種でない点が論拠として脆弱である。”FanGraphs” によると、2002 年以降から MLB 全体でのスプリッター（表内 SF% ）の使用割合は 2%を上回ったことが無く、これはナックルボール（KN%）、球種

³ Wilson (2014)

⁴ The New York Times (2011)

不明（XX%）を除く球種で最低割合である（表-3 参照）。なお、表内項目の詳細については後に示している表-4 を参照頂きたい。

表-3 2002～2015 年の MLB 全体における各球種投球割合の推移

Season	FB%	SL%	CT%	CB%	CH%	SF%	KN%	XX%
2002	0.644	0.121	—	0.112	0.098	0.018	0.008	0.07
2003	0.638	0.118	—	0.1	0.12	0.017	0.007	0.059
2004	0.626	0.136	0.01	0.091	0.114	0.017	0.006	0.089
2005	0.619	0.148	0.022	0.084	0.109	0.014	0.005	0.055
2006	0.611	0.151	0.028	0.088	0.106	0.014	0.003	0.027
2007	0.606	0.149	0.029	0.087	0.113	0.014	0.004	0.021
2008	0.607	0.156	0.036	0.083	0.101	0.012	0.005	0.02
2009	0.597	0.149	0.046	0.091	0.1	0.013	0.004	0.019
2010	0.587	0.148	0.047	0.09	0.107	0.014	0.007	0.016
2011	0.578	0.146	0.057	0.094	0.104	0.015	0.006	0.008
2012	0.576	0.144	0.058	0.104	0.098	0.017	0.004	0.006
2013	0.578	0.145	0.057	0.097	0.102	0.017	0.005	0.006
2014	0.577	0.137	0.062	0.099	0.104	0.016	0.005	0.005
2015	0.578	0.146	0.058	0.089	0.108	0.015	0.006	0.014

この期間に MLB でプレーしたプレーヤーはのべ 9088 人いるが、そのうち 1 回でもスプリッターを投げたプレーヤーは 1504 人と、スプリッターの使用者は少ない。MLB 全体への TJS 波及要因として捉える上で、この規模はあまりに限定的である。

2.1.2.2 カッター

カッターは次ページ図-5 のグリップから投じられ、ストレートとスライダーのハイブリッドとなる球種である。2011 年 6 月 11 日に Chen が米スポーツメディア “Sports Illustrated” にアップしたコラム “This Is The Game Changer” にはカッターを “The pitch screams toward the hitter with the speed and the spin of a fastball and on a plane as flat as a vinyl LP and then, just as it begins to cross the plate, the ball darts like a

badminton birdie.”⁵ と記しており、ストレートの亜種と捉えられるケースが MLB においては主な見立てと言える。

図-5 カッターのグリップ



参照：Wikipedia－カット・ファスト・ボール,

<https://ja.wikipedia.org/wiki/%E3%82%AB%E3%83%83%E3%83%88%E3%83%BB%E3%83%95%E3%82%A1%E3%82%B9%E3%83%88%E3%83%BB%E3%83%9C%E3%83%BC%E3%83%AB>

カッターは上表-3 では 2002 年のデータ集計以降顕著な使用割合の上昇を辿っており、2004 年から 2014～2015 年にかけてその割合は約 6 倍を示している。使用割合で大きな変化が見られない点が根拠として脆弱と捉えられたスプリッターに対し、カッターはこの点で TJS と一定の相関が見受けられる側面から支持を得ている仮説だ。ボルティモア・オリオールズはこの説を支持しているクラブチームであり、2012 年に Dan Duquette が GM として入団してからはチーム全体の方針としてカッターを禁止している。

一方で、カッターを根拠とする上では動力学上に要因を見出す必要がある。オリオールズ投手育成部門のトップである Rick Peterson は Chen (2011) にて、カッター使用によるリスクについて以下の言及を行っている。

“What happens is you start to get off to the side of the baseball (with your grip) and then you’re no longer consistently behind the baseball. Typically what we see is the more you throw that cutter, you can become dependent on it and you start to overuse

⁵ Albert Chen (2011)

it and typically what happens to guys that overuse the cutter is their fastball velocity drops. That has been consistent over the years.”

It’s the nature of the mechanics of throwing that pitch. Power guys that throw the fastball stay behind the ball consistently. When you start throwing that cutter you start getting off to the side of the baseball. If you want to take velocity off a pitch, get off to the side of the baseball. Your intent to throw the cutter is to take velocity off the fastball and flatten it out.”⁶

ここで Peterson はカッター使用による身体能力、及びそれに付随する形でのストレートの球速が低下する点について言及をしており、UCL に与える負担の増大、もしくはそれに影響を与える内容については触れていない。スプリッター同様に、カッターも故障を生み出す特定要因と捉える上では脆弱さが拭えない仮説に止まっている。

2.1.3 メカニクス上の欠陥

ここにおける「メカニクス上の欠陥」とは、下図-6 に示される “Inverted W” と称されるメカニクスのことを示す。Inverted W とはメカニクスにおいて、足の着地時点で両腕のヒジが肩よりも上に位置する形から投球を行うタイプのものを指し、球速の増大に一定の効果をもたらす一方で投手のヒジ及び肩に重大な負担をもたらす動作であると認知されている。Inverted W が投手に与える負担について、Lindsay Berra は 2012 年 3 月に ESPN “Force of habit” にて以下のように整理している。

“（中略） a sign that his sequence is off and he's fighting his own body. Such poor timing leads to arm lag, evident when the throwing elbow trails the shoulder once the shoulders square to home plate. Strasburg（注：MLB・ワシントン・ナショナルズでプレーする Stephen Strasburg のこと） exhibits both problems, forcing him and others like him to rely more on the arm's relatively small muscles instead of the more massive ones in the legs and torso. Throw after throw, the shoulder and elbow

⁶ Chen (2011)

are under extra stress. The higher the pitch's velocity and the worse the flaw, the more the arm suffers. And the more a pitcher throws, the worse it gets.”⁷

すなわち、ヒジを肩よりも上げる動作を入れることで投球時に下半身と上体のタイミングにズレが生じ、腕が遅れて出てくる結果として球速が上昇すると Berra は整理している。その一方で Inverted W はこのようなタイミングのズレから投球時の負担を下半身ではなく、腕の小さな筋肉に依拠する性質を有しており、この点から TJS の発生要因として検討されている。このような側面上、上記仮説にてストレートの球速向上を TJS 要因として支持している ASMI も間接的に Inverted W を TJS 発生要因として支持している形となる。

メカニクス上の欠陥を要因として認識する上での課題は、Inverted W に代表される劣悪なメカニクスが必然的な故障を生み出す訳では無い点にある。すなわち、次に仮説として紹介する「身体的差異」も含めての負担の度合を検討する必要がある、また数値データにより示されない分野に依拠するこの点の解消は実証分析の性質上困難である。

図-6 C.J. Wilson に見られる Inverted W



参照 : Chris O'Leary, "Death To The Inverted W",
<http://www.chrisoleary.com/projects/PitchingMechanics101/Essays/DeathToTheInvertedW.html>,
ChrisO'Leary.com, アクセス日 : 2015 年 11 月 29 日.

⁷ Berra (2012)

2.1.4 身体的差異

MLB においては “There is no such thing as a Pitching Prospect” たる価値観がその略称 “TINSTAAPP” たる形で浸透している。この 1 文は「ピッチングプロスペクト（＝投手の有望株）たるものは存在しない」との意味であり、米スポーツ系シンクタンク “Baseball Prospectus” の設立者である Gary Huckabay により作られた造語である。

米メディア “Yahoo! Sports” のライターである Jeff Passan は MLB 誌『月刊 Slugger』2009 年 8 月号内 “Stephen Strasburg – Million Dollar Baby” にて、TINSTAAPP の本質が投手のコンディション維持の困難さにあることを言及している。Passan は投手各人には「ゴムのように腕が柔らかい投手もいれば、バルサ材でできたような硬い腕の投手もいる」⁸とコメントしており、この部分の予見が出来ない点が TINSTAAPP たる状況を生み出しているとの見立てを示している。上記した各要因と TJS 発生に非常に頑健な相関が見受けられた場合でも、その要因が TJS 発生の必要条件になるとは限らない。この時、そのような不確実性を生み出す要因として、投手各人の身体的差異は大きな影響を及ぼしている可能性について検討することが出来る。

統計学を利用した要因分析を行う上で、これはメカニクスと同じく主体別に内容が変化する部分については本来固定効果として抽出を行うべき部分である。しかし、本論では身体的差異が TJS 発生要因としてどれだけの強弱を有しているかを数値として検討することが出来ない点から、メカニクスと同様に固定効果を内包するモデルの利用は難しいと捉えるべき要因である。

2.2 先行論文

次に、故障要因の実証分析を行った先行論文として、Andrews et al.(2006)の分析アプローチを参照する。

2.2.1 Andrews et al. (2006) における実証分析アプローチ

Andrews et al.(2006)の論文は、1994～1999 年から 2000～2005 年の区分で比較をした際、ヒジの故障例が高校生が 6 倍、大学生で 4 倍になっているデータを根本的な問題意識

⁸ Passan (2009)

の起点としている。そのような故障発生の要因分析の手法として、この論文では t 検定、及びロジットモデルの推定を行い、実証分析による故障要因の特定を試みている。

2.2.1.1 t 検定・ χ^2 検定

本論で参照したい実証分析のアプローチを Andrews et al.(2006) は以下の方式を採用している。まず t 検定について、Andrews et al.(2006, $p.906$)では以下のような分析上の目的意識が記述がなされている。

“Although this small study provided an informative description of young pitchers who required surgery, it did not compare those pitchers with a healthy control group to determine risk factors. The current study attempts to make that comparison.

The purpose of this study was to identify risk factors that could predispose an adolescent pitcher to a significant shoulder or elbow injury. We hypothesized that there are significant differences in pitching practices between the healthiest adolescent pitchers (ie, pitchers with no history of problematic shoulder or elbow pain) and the most seriously injured adolescent pitchers (i.e., pitchers who have required either shoulder or elbow surgery for a pitching-related injury).”

ここでは先行論文として参照している Lyman et al.(2001)、Petty et al.(2004)、USA Baseball (=アメリカ野球協会) に対し、コントロールグループの設定たる視点が抜けているため、故障プレーヤーと健常プレーヤー間における差異に関しての言及が欠如しているとの問題点を示している。Andrews et al.(2006) はこの点の解消に向けて、基本統計量として得られた全 140 人の主体を「肩を故障したグループ」(=Shoulder)、「ヒジを故障したグループ」(=Elbow)、「故障グループ全体」(=Injured)、「健康なグループ」(=Control) の 4 グループに区分、その上で Control とその他グループ間における各データ項目の平均の差異を有意水準 1%とした t 検定により検証するアプローチを採用している。ただし、この論文では基本統計量の収集をアンケートにより行っている関係から、離散型の変数を得るデータも存在する。そのようなデータの実証分析に関しては χ^2 検定を採用している。ま

た、これらの統計学的な理論背景に関しては次章で詳細を説明する。

2.2.1.2 ロジットモデル

ロジットモデルについては、モデル内に設定する変数の決定方法として Andrews et al.(2006, p.907) では以下の記述がなされている。また、ロジットモデルの統計学上の理論背景についても、 t 検定同様次章で説明を行う。

“A multivariable logistic regression model was developed to identify variables independently associated with the injured group. Variables considered for the model were selected based on the results of univariate analyses ($P < .1$). Models were developed with forward stepwise (conditional) regression. Variables were retained in the model if the Wald test had a P value less than .05. Each continuous variable identified by the multivariable logistic regression model was then divided into a high group and a low group, and odds ratios were determined.”

ここではステップワイズ法による変数の抽出プロセスとして、まず全てのデータ項目に対して単一変数によるロジットモデルの推定を行った後、設定した有意水準で係数が有意になる変数による多変数のロジスティックモデルを構築し、さらに Wald 検定による変数抽出を実施している。単一変数ロジットモデルでは 10%、その後に行う Wald 検定では 5%を有意水準として設定して行っている。

2.2.2 Andrews et al.(2006) の分析結果

以上の実証分析アプローチを用いた結果を以下に整理する。

2.2.2.1 t 検定・ χ^2 検定

Andrews et al では t 検定・ χ^2 検定の結果として、Injured で 16、Elbow は 15、Shoulder は 10 のデータ項目に関して有意な差異を示しており、論文内でもこれらが要因として有力な候補になると捉えている。特に分析結果を踏まえての言及では、Lyman et

al (2001)や Petty et al (2004) 等の先行研究では故障要因として捉えられていた「変化球の投球割合 (=Out of 10 pitches, how many were)」が有意では無いとの結果を得た理由について取り上げており、いずれのグループにおいても変化球の習得・利用が早い年齢から行われている可能性を示唆している。

2.2.2.2 ロジットモデル

ロジットモデルについては下図-7 に整理する。

図-7 Andrews et al (2006) におけるロジットモデルの結果

Variable	Odds Ratio	95% Confidence Interval
Months per year of competitive pitching		
≤8	Referent	
>8	5.05	1.39-18.32
No. of pitches per appearance		
≤80	Referent	
>80	3.83	1.36-10.77
Fastball speed, mph		
≤85	Referent	
>85	2.58	0.94-7.02
How often pitch with arm fatigued		
Never	Referent	
Infrequently	4.04	0.97-16.74
Regularly	36.18	5.92-221.22

^aGoodness-of-fit (Hosmer and Lemeshow test) = 0.755.

この結果を参照すると、連続変数においては特定の値をボーダーとする形で 0・1 のダミー変数への変形を行い、離散変数と同じ形式での処理が出来るようにデータを加工している。ただし、このモデルを設定する前に行った単一変数によるロジットモデルについては論文内で結果の明示が無く、その際にも連続変数データを同様に扱ったかについては定かでは無い。

ここでは計 4 項目が説明変数として選択がされている。また、図内 “Odds Ratio” と示さ

れている部分は故障リスクの上昇度合を示しており、各データ項目の結果により与えられるリスク度合の計量化に成功している。

2.2.3 Andrews et al (2006) から本論が参照すべき点

先行論文から参照すべき点は以下 2 点である。

- TJS の発生要因を分析する上で、その被説明変数は TJS が発生するか否かを示す 0・1 での確率である。従って、本論においても確率の推定モデルとしてロジットモデルを利用する必要がある、そのモデル設定のアプローチについて Andrews et al. (2006) を参照すべきである。また、グループ分けを行った上での差異分析に関しても同様である。
- 一方で、Andrews et al. (2006)におけるロジットモデルの推定はダミー変数化のプロセスに一定の恣意的な設定が見受けられる。論文内においてはダミー変数化のボーダー設定に対し理論的な理由付けが無く、元々連続変数であるものに対して行う処理としては適切では無いと考える。Andrews et al. (2006) 自身も言及部分では基本統計量をアンケート収集に頼った点から精度が担保出来ていないとの課題を記しており、この点から本論では基本統計量として連続変数のみを使用する形を模索するアプローチを採用する。

3. 基本統計量と実証分析手法

本章では本論で行う TJS 要因についての基本統計量、及び実証分析にて使用する分析手法について整理する。

3.1 基本統計量

本論では FanGraphs.com (<http://fangraphs.com>) を使用して 2002～2015 年の期間にて MLB でプレーした全投手のパネルデータ（主体計 2345 人、サンプル数 9102）を作成し、分析に使用した。パネルデータに入るデータ項目は下表-4 に整理する。

表-4 パネルデータに入るデータ項目

データ項目名	意味（単位）
IP	イニング数（IP）
Pitches	投球数（球）
FB%	ストレートの使用割合（%）
SL%	スライダーの使用割合（%）
CT%	カッターの使用割合（%）
CB%	カーブの使用割合（%）
CH%	チェンジアップの使用割合（%）
SF%	スプリッターの使用割合（%）
FBv	ストレートの平均球速（mph）
SLv	スライダーの平均球速（mph）
CTv	カッターの平均球速（mph）
CBv	カーブの平均球速（mph）
CHv	チェンジアップの平均球速（mph）
SFv	スプリッターの平均球速（mph）
TJD	TJS 経験に対し 1 を得るダミー変数

なお、TJD は TJS を経験した年度以前全てのシーズンに 1 を得て、直近の TJS を実施して以降のシーズンには 0 を付ける。これはパネルデータとしてデータを収集した点も含め、TJS の発生前後を時系列的に捉えた上でダミー変数を設定したいとの分析上の意図をもった処理である。第 2 章で示した TJS に関する仮説の存在から、TJS 前後ではメカニクスや使用球種の割合等、プレースタイルが少なからず変様する可能性を検討することが出来る。本論では TJS 前後では同じ選手でも別のプレースタイルを有する別主体として捉えた上で、TJD の設定を行った。なお、TJD が 1 のサンプル数は 860、0 は 8271 である。

また、下表-5 ではパネルデータ全体、及び TJD の 0・1 で区分したグループ別に各データ項目の平均と標準偏差（SD）を示す

表-5 基本統計量

データ	全体平均	全体 SD	TJD"1"平均	TJD"1"SD	TJD"0"平均	TJD"0"SD
IP	65.8330	62.3448	80.4206	64.1383	64.3161	61.9625
Pitches	1076.378	984.322	1307.286	1010.878	1052.378	978.4595
FB%	0.6153	0.1226	0.6386	0.1066	0.6128	0.1239
SL%	0.1892	0.1185	0.1749	0.1068	0.1907	0.1196
CT%	0.1289	0.1461	0.0950	0.1137	0.1324	0.1486
CB%	0.1211	0.0914	0.1191	0.0940	0.1213	0.0910
CH%	0.1097	0.0832	0.1013	0.0813	0.1105	0.0834
SF%	0.0923	0.1146	0.0654	0.0662	0.0954	0.1186
FBv	90.6266	3.0752	91.4323	2.7181	90.5425	3.0982
SLv	82.5118	2.9760	82.9674	2.9882	82.4630	2.9707
CTv	86.8874	2.7441	87.0382	2.9569	86.8715	2.7211
CBv	76.4064	3.3994	77.0359	3.2918	76.3381	3.4041
CHv	82.3509	3.1687	82.7752	2.9570	82.3046	3.1876
SFv	83.7760	4.7561	84.1359	3.0057	83.7336	4.9201

表-5 を参照すると、IP、Pitches、そして球速に関するデータ項目（FBv～SFv）は全て TJD で 1 を得るグループが高い平均値を得た。使用球種の割合についてはストレート（FB%）とスプリッター（SF%）で TJD1 グループが高い値をとる一方、その他の球種に関しては TJD0 グループの使用割合が高くなった。

3.2 分析手法

本論では TJS が発生する要因についての実証分析を目的としており、その実現に向けては「*t*検定」「ロジットモデル」の 2 手法による分析を行う。

3.2.1 *t* 検定

t 検定とは、互いに独立な異なる 2 グループのデータが正規分布に従う仮定に沿い、2 グ

ループの平均値が等しいかを検定する手法である。以下では t 検定式について、いくつかのケースに分けて整理する。

3.2.1.1 等分散における t 検定

標本数 n_1 ・標本平均値 μ_1 と設定されるグループ 1 と、標本数 n_2 ・標本平均値 μ_2 であるグループ 2 の平均の差異を検定する t 検定統計量は

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\beta_e \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

と定義される。この時 β_e とはグループ 1・グループ 2 がいずれも等分散を有した正規分布に従うとの仮定の下での分子の分散の推定量であり、標本分散がグループ 1 で β_1 、グループ 2 で β_2 を得る場合に

$$\beta_e = \frac{(n_1 - 1) \cdot \beta_1 + (n_2 - 1) \cdot \beta_2}{n_1 + n_2 - 2}$$

との加重平均により得る。すなわちこの検定は自由度 $v = n_1 + n_2 - 2$ の t 分布に従っている。検定統計量 t を用い、帰無仮説「 $\mu_1 - \mu_2 = 0$ 」の検定を行う⁹。

3.2.1.2 異分散における t 検定 (Welch 検定)

等分散 β_e を仮定せず、各グループが異分散を有すると設定をした上で 2 グループ間の平均値における t 検定は Welch 検定と呼ばれる。すなわち検定統計量 t_w は

$$t_w = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\beta_1}{n_1} + \frac{\beta_2}{n_2}}}$$

⁹ 花田 (2006)

と定義される。また、この時 t_w は t 分布に従い、その自由度は

$$v = \frac{\left(\frac{\mu_1}{n_1} + \frac{\mu_2}{n_2} \right)^2}{\frac{\beta_1^2}{n_1^2(n_1-1)} + \frac{\beta_2^2}{n_2^2(n_2-1)}}$$

を得る。

3.2.1.3 本論における t 検定の利用

TJS が発生する要因を決定する上では、TJS 経験の有無たる点において投手をグループ分けした場合、TJS 要因となるデータ項目では有意なグループ間の差が確認される必要がある。本論では後述する基本統計量を「TJS 経験を有するグループ」「TJS 経験が無いグループ」に分類した上で、上表-4 に示したデータ項目について上述の検定を行う。 t 検定における分散の決定には F 検定を行い、帰無仮説「2 グループは等分散を有する」について検定分析をする。 F 検定は上述グループ 1・2 のケースを用いると、以下式から F 検定量を得る。

$$F = \frac{\beta_1}{\beta_2} \quad (\text{ただし、} \beta_1 < \beta_2)$$

本論では 2 グループによる t 検定において 5% 有意水準を設定し、そのグループ間差異を検定した。この時、全てのデータ項目でグループ間での有意な差異が認められなかった場合は「統計学上実証出来る TJS 要因は存在しない」との結論を得る。

3.2.2 ロジットモデル

ロジットモデルは、「特定の事象が発生する/しない」等イベントについて示した質的従属変数が被説明変数である場合に用いられる回帰モデルである。

主体 i における事象 x の発生について「1」を得、発生しなかった場合「0」を得る 2 値の

質的従属変数 Y_i とすると、このイベントは

$$Y_i = \begin{cases} 1 & (\text{確率 } P_i) \\ 0 & (\text{確率 } 1 - P_i) \end{cases}$$

となる。この時主体 i において事象 x が発生する確率は P_i だが、その値は主体 i に作用する各要因の属性によって決定される。よって、ここでは P_i を主体 i の特定の関数として設定する。その関数にロジスティック関数を使用するモデルがロジットモデルである。

ロジット変換は Y_i が 1 を得るオッズ、すなわち $\frac{P_i}{1-P_i}$ の対数を得る。その変換値に対する回帰式は

$$\log\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \dots + \beta_n \cdot x_{ni}$$

が設定され、実証分析においてはこの推定式における係数が導かれる。この左辺を P_i に戻した回帰式が

$$P_i = \frac{\exp(\alpha + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \dots + \beta_n \cdot x_{ni})}{1 + \exp(\alpha + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \dots + \beta_n \cdot x_{ni})}$$

となり、これがロジットモデルとなる。ロジットモデルを用いることで、設定された説明変数の変動が 0 から 1 の値を得る確率に与える作用について検討出来る¹⁰。すなわち、各リスク要因に関してその具体的な影響を定量化出来る点にこのモデルは価値を有している。

本論では上述した基本等計量において TJS の経験の有無により 0 と 1 を得るダミー変数 ”TJD” を設定し上記 P_i とした。なお、分析においては取得データの形式上、パネル 2 項ロジットモデルを採用することで、各主体別の固定効果を導出した上でモデルを得るアプローチも検討を行ったが、固定効果に内包される要因・意味を設定するモデル上具体的な定義が難しい点を考慮し採用を見送っている。

¹⁰ 花田 (2006)

3.3 本論における分析の手順

本論では以下の手順で分析を行った。

1. 得られた基本統計量に関して F 検定を行い、各データ項目の分散について等異の定義を行った。
2. F 検定の結果に沿い 5%有意水準で t 検定を適切な設定の下実施し、TJS 経験の有無で平均値に有意な差が生じているかを確認した。
3. ロジットモデルを設定した。説明変数の設定を行う上では Andrews et al. (2006) を参照した方法を参照し、TJD との単一ロジットモデルを全データ項目に対し実施した後、有意水準 10%を満たしたデータのみを選択して AIC 基準によるステップワイズ法によるモデル設定のアプローチを採用した。

4. 分析結果

ここでは設定した分析アプローチに沿い、実証分析を行った結果を整理する。

4.1 t 検定

ここでは、TJS 経験の有無によるグループ分けで各データ項目に有意な差が生じるかの検定について結果を整理する。

まず、 t 検定において定める分散の性質について整理する。以下表-6 は、上表-4 に示したデータ項目において F 検定を行った結果を示している。前章で述べた通り、ここでは有意水準を 5% に設定し「2 グループの分散は等しい」たる帰無仮説を検定している。

表-6 各データ項目における F 検定の結果

データ項目名	p-value	2 グループの分散設定
IP	0.1652	等分散
Pitches	0.1899	等分散
FB%	$5.481 \times 10^{-10}***$	異分散
SL%	$1.34 \times 10^{-5}***$	異分散
CT%	$2.2 \times 10^{-16}***$	異分散
CB%	0.4348	等分散
CH%	0.115	等分散
SF%	$2.2 \times 10^{-16}***$	異分散
FBv	$2.2 \times 10^{-16}***$	異分散
SLv	0.03686*	異分散
CTv	0.807	等分散
CBv	0.5201	等分散
CHv	$1.381 \times 10^{-9}***$	異分散
SFv	0.03308	異分散

(注) *** : 0.1%有意 ** : 1%有意 * : 5%有意 . : 10%有意

以上を用いた上で、 t 検定を行った結果を次ページ表-6 に示す。

表-7 t 検定

データ項目名	p-value
IP	$5.224 \times 10^{-13}^*$
Pitches	$4.563 \times 10^{-13}^*$
FB%	$4.563 \times 10^{-13}^*$
SL%	0.05322*
CT%	0.001017*
CB%	0.5496
CH%	0.1654
SF%	0.01747*
FBv	$2.2 \times 10^{-16}^*$
SLv	0.02524*
CTv	0.8629
CBv	0.03474*
CHv	$1.043 \times 10^{-5}^*$
SFv	0.1045

上表-7 において “*” 付のデータ項目が 5%有意水準を満たしている。この結果を踏まえると、前章の表-5 に示した平均値の差について、以下の内容を言及することが出来る。

- TJD1 グループが TJD0 グループの平均を上回ったデータ項目のうち、有意な差を統計学的に検証が出来たものは「IP」「Pitches」「FB%」「SF%」「FBv」「SLv」「CBv」「CHv」の 8 項目であった。すなわち、これらの球種において TJD1 グループはよりハードに投げており、かつストレート・スプリッターに関しては積極的に使用していることが示されている。また、プレーする絶対量から見ても、TJS を経験していないグループよりも献身的であることが伺える。

- 一方でグループ間の差異が有意でないとの結果が得られたのは「SL%」「CB%」「CH%」「CTv」「SFv」であり、TJS 経験の有無がこれら球種の使用割合、及び球速に影響を与える要因では無いとされた。これは TJS の防止策としてこれらの球種の使用を増やすこと以上に、ストレート・スプリッターの使用を抑えることが意味が大きいとの示唆ととることが出来る。

4.2 ロジットモデル

リスク要因の定量化を目的としたロジットモデルについても結果を整理する。

4.2.1 単一変数によるロジットモデル

まず、説明変数の選択を目的とした単一変数によるロジットモデルを検証し、10%有意水準を満たすデータ項目を抽出した。次ページ表-8 では単一ロジットモデルの推定結果を示している。

表-8 単ーロジットモデルの推定結果

データ項目名	Intercept	データ項目係数
IP	-2.35420***	0.00376***
Pitches	-2.545***	2.396×10^{-4} ***
FB%	-3.3821***	1.7889***
SL%	-2.0155***	-1.1769***
CT%	-2.00251***	-2.18078***
CB%	-2.18800***	-0.27332
CH%	-2.06901***	-1.43248**
SF%	-1.8755***	-3.3264**
FBv	-11.86217***	0.10551***
SLv	-7.12591***	0.05915***
CTv	-4.21360.	0.02253
CBv	-7.08326***	0.06339***
CHv	-6.24751***	0.04883***
SFv	-4.48854.	0.02796

(注) *** : 0.1%有意 ** : 1%有意 * : 5%有意 . : 10%有意

本論では前章末で示した手順に沿い、「CB%」「CTv」「SFv」の3項目以外のデータを利用してのステップワイズ法を行うこととした。また、表-8 よりステップワイズ法のプロセスから除外された3項目はいずれも t 検定においてグループ間の有意差が存在しないとの結果と一致するものであり、2種の統計学的アプローチにおいて TJS 要因として検討すべき余地が無いことを示している。

4.2.2 ステップワイズ法により採択されたロジットモデル

前節で選択された説明変数を用い、ステップワイズ法による AIC 基準でのロジットモデルの設定を行った。最小となる AIC を得るロジットモデルとしては、以下表-9 を得ることが出来た。

表-9 ステップワイズ法により選択された最小 AIC ロジットモデル

データ項目名	Estimate	p-value
Intercept	-12.36	$2 \times 10^{-16}***$
FBv	9.228×10^{-2}	$1.08 \times 10^{-11}***$
IP	4.061×10^{-3}	$1.59 \times 10^{-13}***$
FB%	1.763	$3.17 \times 10^{-7}***$
CHv	4.18×10^{-3}	$0.00545**$
AIC : 5548.3		

これらの説明変数の変動が発生確率 P_i に与える影響は、上表のように推定されたロジットモデルの立式

$$P_i = \frac{\exp(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n \cdot x_{ni})}{1 + \exp(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n \cdot x_{ni})}$$

から捉えると、例えば説明変数 x_{1i} の微小な変化に対する P_i の変化については

$$\frac{\partial P_i}{\partial x_{1i}} = \frac{\beta_1 \exp(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n \cdot x_{ni})}{(1 + \exp(\alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n \cdot x_{ni}))^2} = \beta_1 \cdot P_i \cdot (1 - P_i)$$

であり、これは変動が与える影響が全ての説明変数 $x_{ki|k=1,2,\dots,n}$ に依存して決定されることを示している。次ページ表-10 では、表-5 に示される標本平均での P_i を基準とした際の各変数の微小な変化に対する P_i の変化量を整理する。

表-10 標本平均を基準とした際の説明変数の微小な変化に対する P_i の変化量

データ項目名	P_i の変化量
FBv	7.639×10^{-3}
IP	3.361×10^{-4}
FB%	0.1459
CHv	3.460×10^{-4}

また、下表-11 は表-9 で得られたモデルにより得られたパネルデータにおける TJS 発生確率の推定結果である。

表-11 ロジットモデルによる TJS 発生確率の推定結果

年度	選手名	TJS 発生確率
2009	Justin Verlander	0.26604
2011	Justin Masterson	0.25331
2003	Jason Schmidt	0.24969
...		
2003	Jarred Fernandez	0.007621
2005	Tim Wakefield	0.006889
2010	Bryan Shaw	5.40×10^{-6}

ロジットモデルにて以上 4 変数が選択された理由については、以下を検討することが出来る。

- FBv・FB%

この 2 点は仮説「MLB における球速向上」と合致し、ストレートをより多く、よりハードに投げることが TJS 要因と成り得ることを示している。いずれも IP、CH%と比較し TJS 発生に与える影響は大きく、それについては上図-10 から明らかである。ハードなストレートを投じることが Inverted W となるメカニクスを引き起こしていると複合的な要因が入った結果である可能性も検討出来るが、本論では処理を行ったデータの関係上その点に

関する言及は差し控える。

● IP

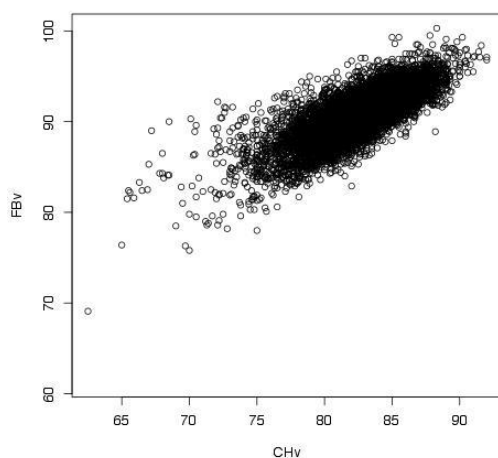
IP が要因として入ったことは、絶対的な投球量の増大も TJS 発生において影響を及ぼしていると解釈が出来る。すなわち、ストレートの割合、球速がいずれも MLB 平均を下回る水準であった場合でも、年度を辿り投球量が蓄積されることで TJS を発生させる可能性を検討することが出来る。

一方で、本論ではデータ項目として投球数 (Pitches) も入れた上での処理も行ったが、こちらに関してはステップワイズ法での処理時にモデルから選択をされなかった。投球量として Pitches では無く IP が選択をされた点は、同じ投球数をクリアした場合でもより長い IP を消化する可能性が大きい先発タイプの方が TJS を発生させやすいとの言及を得ることが出来る。

● CHv

CHv は下図・9 で示すように、FBv と一定の正の相関を有している。すなわち CHv とは FBv を補完する意味で選択された変数であり、球速の増大が TJS 発生に与える影響として非常に強いことを示している。

図-9 FBv と CHv の相関プロット



5. インプリケーション・課題

本論では TJS 発生に関する仮説、及び故障要因の実証分析アプローチを ASMI の先行論文から参照した後、MLB のパネルデータを利用した t 検定とロジットモデルにより TJS 発生要因について以下の言及を得た。

- t 検定により TJS 経験の有無で統計学上有意な差異を有するデータ項目としては「IP」「Pitches」「FB%」「SF%」「FBv」「SLv」「CBv」「CHv」を示すことが出来た。これらはいずれも TJS 経験を有する P が TJS 経験の無い P を上回る値を取っている。
- また、本論ではロジットモデルによるリスク要因の定量化も行い、ステップワイズ法により選択された回帰モデルの説明変数は「FBv」「IP」「FB%」「CHv」となった。これらは t 検定で得られた結果をより厳密に絞り込んだものとして捉えることが出来、TJS 発生リスクの予見において特に重要な要因になると本論では結論出来る。また、その要因については前章で述べた通りであり、そちらを参照されたい。

また、本論における分析上の課題としては以下の 2 点を取り上げる。

1 点目は、本論で使用したパネルデータ収集についての問題点である。本論ではパネルデータの収集を 2002～2015 年の全投手に対し行ったが、その理由は時系列的に TJS 前後のパフォーマンスを整理したいとの意図を持っての処理であった。しかし、このパネルデータ自体は特に割合では無く絶対数を示す Pitches と IP で大きな分散が見受けられ、分析を行う上で決して無視することが出来ない条件面の非統一性を招いているとの課題が存在する。本論が分析において表現を試みたデータの連続性とパネルデータそのものの質を両立することは困難であり、このバランスをいかに整えるかが分析上の課題となる。

また、2 点目はロジットモデルで 2 項パネルロジスティックモデルを採用しなかった点である。本論では分析の目的上、主体別で差異が存在する部分については固定効果たる形として抽出を行った上で回帰モデルを設定するのが最も理論上適切な処理であった。しかし、本論では抽出すべき主体別の差異が仮説における「メカニクス上の欠陥」「身体的差異」であり、これらの数値データとしての表現が出来ない点からこのモデルの使用を断念している。実証分析精度の向上を目指す上ではこの部分の解消が必須であり、特に身体的差異に関し

ては身体測定データから身体強度を推定する等、何らかの手段を講じる必要がある。

TJS 発生要因に関して複数の仮説が提唱される昨今において、実証分析の見地から一定の結論を得ることが出来た点は本論の貢献である。上記課題の解消と共に、TJS 発生に関するブラックボックスがより明らかとなることを願いつつ、本論の結びとさせて頂く。

参考文献・データ

1. 出野哲也 (2014) 「ヒジの故障は何が原因で起こるのか？」月刊 Slugger 2014 年 10 月号、PP14-17.
2. 花田憲三 (2006) 「実践的多変量解析法」、日科技連出版社、pp.122-130.
3. Andrews. J, Olsen. S, Fleisig. G, Dun. S, Loftice. J, (2006) "Risk Factors for Shoulder and Elbow Injury in adolescent baseball pitcher", The American Journal of Sports Medicine, pp.905-912.
4. Lyman S, Fleisig GS, Waterbor JW, et al. (2001) Longitudinal study of elbow and shoulder pain in youth baseball pitchers. *Med Sci Sports Exerc*, 33, pp.1803-1810.
5. Petty DH, Andrews JR, Fleisig GS, Cain EL (2004). Ulnar collateral ligament reconstruction in high school baseball players: clinical results and injury risk factors. *Am J Sports Med.*, 32, pp.1158-1164.
6. Passan. J (2009) "Stephen Strasburg – Million Dollar Baby" 月刊 Slugger 2009 年 8 月号、pp.18-21.
7. Wilson. J (2014) 「経験者が語るトミー・ジョン手術」月刊 Slugger 2014 年 10 月号、pp. 21-23.
8. ASMI (2014) "Position Statement for Tommy John Injuries in Baseball Pitchers", <http://www.asmi.org/research.php?page=research§ion=TJpositionstatement>,
アクセス日：2015 年 8 月 11 日.
9. Berra. L (2012) "Force of habit",
http://espn.go.com/mlb/story/_/id/7712916/tommy-john-surgery-keeps-pitchers-game-address-underlying-biomechanical-flaw-espn-magazine, アクセス日：2015 年 8 月 25 日.
10. Chen. A (2011) "This Is The Game Changer",
<http://www.si.com/vault/2011/06/13/106077823/this-is-the-game-changer>, Sports Illustrated, アクセス日：2015 年 9 月 12 日.
11. Leaderboard, "FanFrapths", <http://fangraphs.com>, アクセス日：2015 年 11 月 29 日.
12. The New York Times (2011) , "Split-Finger Fastball, Once Popular, Is Falling Away", <http://www.nytimes.com/2011/10/02/sports/baseball/split-finger-fastball-use-of-a-popular-pitch-falls-off-the-table.html>, アクセス日：2015 年 8 月 23 日
13. Tommy John Surgery List (@MLBPlayerAnalys),

<https://docs.google.com/spreadsheets/d/1gQujXQQGOVNaiuwSN680Hq-FDVSCwvN-3AazykOBON0/edit#gid=0>, アクセス日：2015 年 8 月 10 日