

プロ野球の得点分析

ガンマ・ポアソン分布の適用例

松尾 精彦

概要

この論文は、プロ野球の一試合あたりの得点分布がどのような形状をしているかを、2012年シーズンについて調べたことがきっかけである。試しにあるチームの得点分布をガンマポアソン分布にあてはめてみたところ、適合度は極めて良好であった。そこで、セントラル・リーグ（以下セリーグと略記する）、パシフィック・リーグ（以下パリーグと略記する）合わせて12チームについて5シリーズにわたってデータ入力し、統計分析ソフトウェア JMP®を用いて、ガンマ・ポアソン分布に従っているかについて調査したものである。調査してさらに分かったことは、失点の分布もまたガンマ・ポアソン分布に従っているといえることである。延べ60チームの得点分布と失点分布をグラフィカルに表し、適合度の高さをしめす。このあてはまりの良さをを用いれば、プロ野球のレギュラー・シーズンのシミュレーションが可能になる。つまり、この論文は今後のシミュレーションを行うための根拠となる。

1 はじめに

プロ野球の一試合あたりの得点分布が、ガンマ・ポアソン分布に従っていないとは言えないことは、得点データを入力してその棒グラフを表示し、そのグラフがガンマポアソン分布に従っているかを検定したことより判明した。この一連の作業は、統計分析ソフトウェア JMP® を用いて簡単に行うことができた。

そもそも一試合の得点は、通常9イニングの得点の和である。各イニングの得点とは、3アウトになるまでに何人の走者が本塁を駆け抜けたかで決まる。それゆえポアソン分布に従うとは考えづらい。従って一試合の得点分布はポアソン分布に従うとは思えない。また、ガンマ・ポアソン分布に従っているという理論的な説明も難しい。

しかし、実際に分布を当てはめると、一試合の得点の分布は、ガンマ・ポアソン分布により良好に近似されることがデータによって示された。そこで、日本プロ野球連盟主催の2012年度から2016年までの5シーズンを取り上げ、実際の近似の良さを眺めることにする。

この論文では、セリーグ、パリーグ合計12チームの5年間のデータを図表化している。図の見方は左から、得点・失点分布の棒グラフとガンマ・ポアソン分布のあてはめ太線、次に要約統計量、そして右端上に、ガンマ・ポアソン分布を当てはめたときのパラメータの最尤推定値とその信頼限界、その下にピアソン・カイ2乗検定の結果が出力されている。具体的には p 値が表示され、あてはまりの良さが数値的に示される。

第2節では、ガンマ・ポアソンモデルの数学的な定義とパラメータの解釈をおこない、第3節では2012年シーズン、第4節では2013シーズン、第5節では2014シーズン、第6節では2015シーズン、第7節では2016年シーズンを扱う。

2 ガンマ・ポアソン分布の定義とパラメータ

ガンマ・ポアソン分布とは、ポアソン分布、

$$p(y) = \frac{\mu^y}{y!} e^{-\mu}$$

のパラメータ μ が、ガンマ分布、

$$g(\mu; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \mu^{\alpha-1} e^{-\beta\mu}$$

に従うと仮定して得られるもので、両式を掛け合わせ μ について積分すると、ガンマ・ポアソン分布、

$$gp(y; \alpha, \beta) = \frac{\Gamma(y + \alpha)}{\Gamma(\alpha)y!} \left(\frac{\beta}{1 + \beta}\right)^\alpha \left(\frac{1}{1 + \beta}\right)^y$$

が得られる。この分布の期待値と分散は、それぞれ、

$$E(Y) = \frac{\alpha}{\beta}, \quad V(Y) = \frac{\alpha}{\beta} \left(1 + \frac{1}{\beta}\right)$$

となる。上の式は、平均パラメータ $\lambda = \frac{\alpha}{\beta}$ と過分散パラメータ $\sigma = 1 + \frac{1}{\beta}$ を用いて書き換えることができる。最尤法では、 α, β について最大化し、その値から上の関係を利用して λ と σ を求めているのである。

3 2012年シーズン

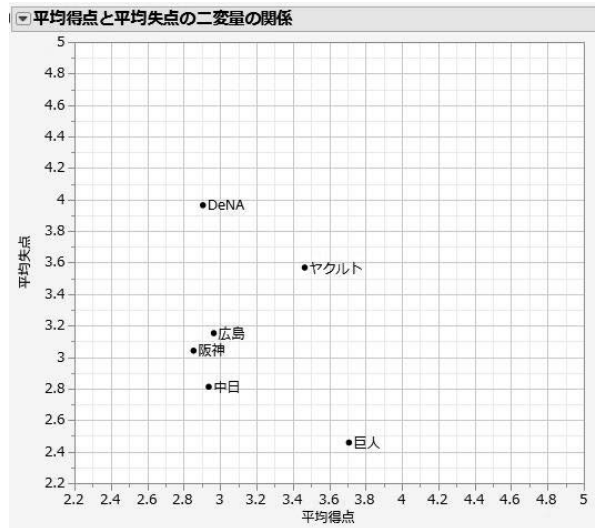
3.1 セリーグ2012年

図2から図7を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

また、表1にもあるように、首位と2位とのゲーム差が10.5、首位と最下位とのゲーム差が31.5と、極めて大味なものであった。各球団の平均得失点を散布図にすると、図1ようになる。この散布図では、点が右にあれば平均得点が高く、点が上にあれば平均失点が高い。よって右下に位置するチームが強かったチームで、左上に位置するチームは弱かったチームであることが分かる。

この散布図で分かることは、巨人が他の5チームから飛び離れた強いチームであると言え、ヤクルトは打撃力はあるが失点も多いチーム、阪神・中日・広島はよく似た投打のバランスを示している。残るDeNAは、得点平均こそ阪神・中日・広島と同等であるが、平均失点が非常に高く、最下位であることがよく分かる。

図1 セリーグ 2012年シーズン平均得失点散布図



なお、図1、図8、図15、図22、図29、図36、図43、図50、図57、図64は、すべて同じ軸範囲で描かれているため、年度毎やセリーグとパリーグの比較などが容易であることに注目されたい。

表1 セリーグ・ペナントレース結果 (2012)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①巨人	144	86	43	15	.667	-
②中日	144	75	53	16	.586	10.5
③ヤクルト	144	68	65	11	.511	20.0
④広島	144	61	71	12	.462	26.5
⑤阪神	144	55	75	14	.423	31.5
⑥DeNA	144	46	85	13	.351	41.0

図2 1位:巨人2012年シーズン

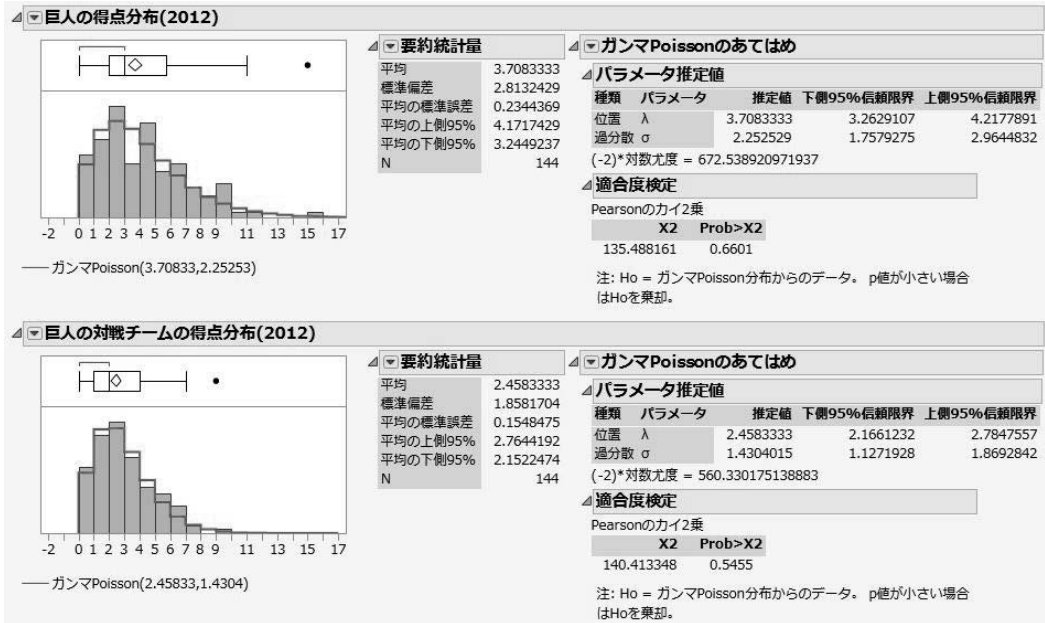


図3 2位:中日2012年シーズン

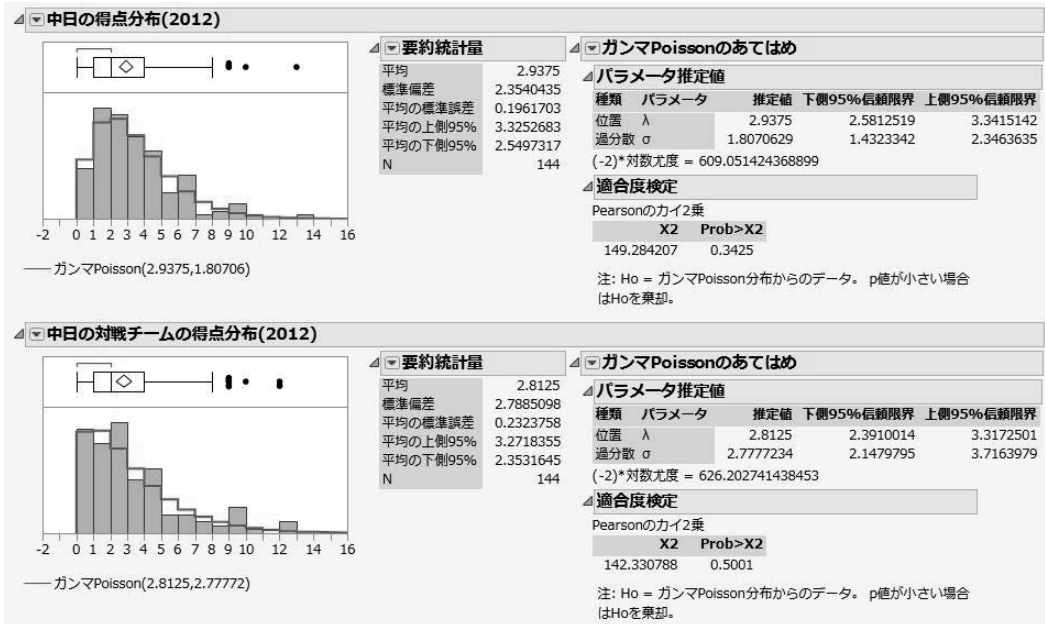


図4 3位:ヤクルト2012年シーズン

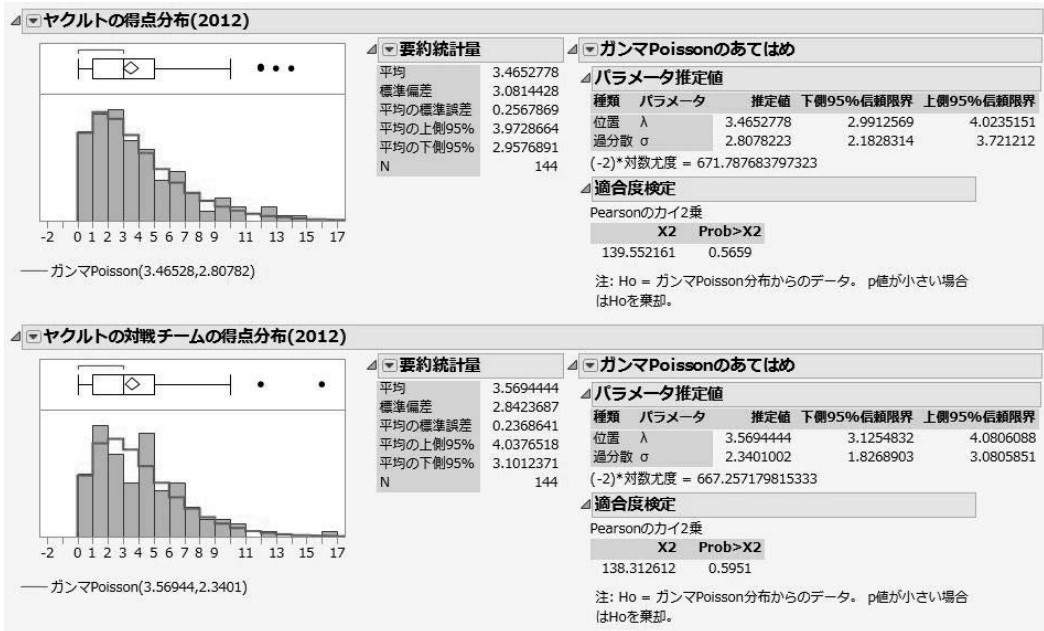


図5 4位:広島2012年シーズン

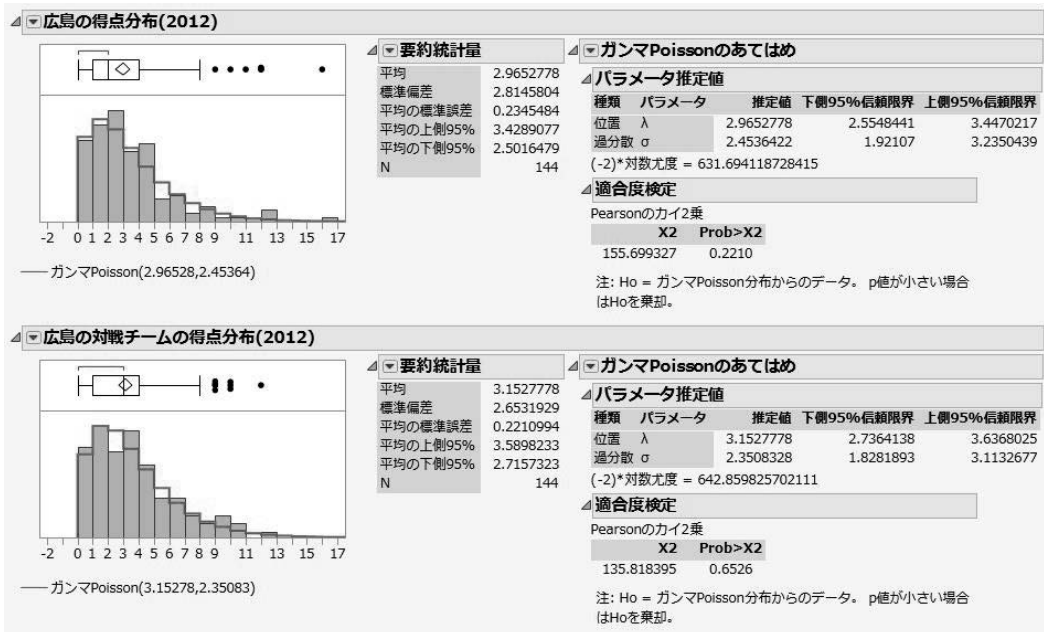
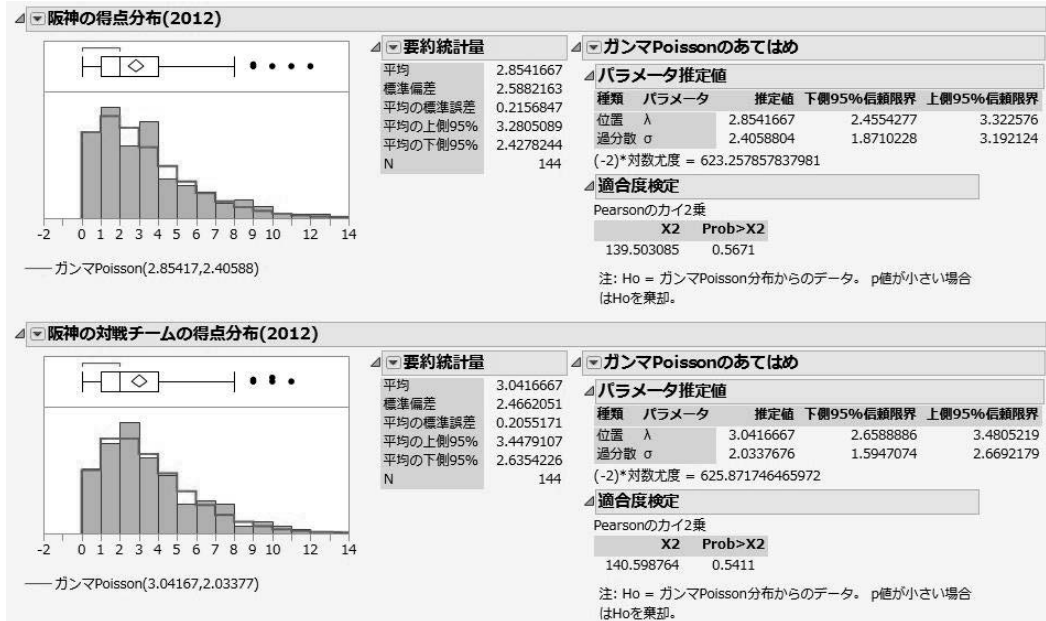


図6 5位:阪神2012年シーズン



3.2 パリーグ 2012 年

図9から図14を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

この年のパリーグは、図8を見て分かるように、散布図の点が互いに近く並んでいて、セリーグに比べて接戦であることが分かる。その中で頭一つ飛び出した日本ハムの優勝で終わった。西武・ロッテ・楽天・ソフトバンクは、得点と失点が正比例するかのように並んでいるのが分かる。残るオリックスは全体のグループから離れ、平均得点は低く、平均失点が高く最下位という結果に終わっている。

表2 パリーグ・ペナントレース結果 (2012)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①日本ハム	144	74	59	11	.556	-
②西武	144	72	63	9	.533	3.0
③ソフトバンク	144	67	65	12	.508	6.5
④楽天	144	67	67	10	.500	7.5
⑤ロッテ	144	62	67	15	.481	10.0
⑥オリックス	144	57	77	10	.425	17.5

図8 パリーグ 2012年シーズン平均得失点散布図

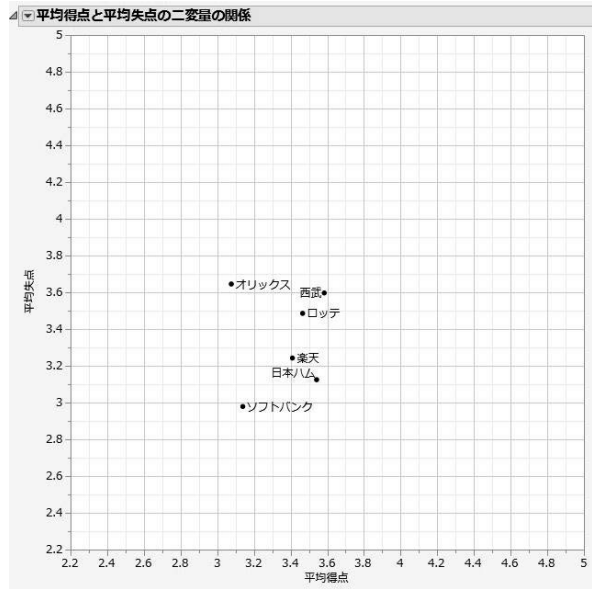


図9 1位:日本ハム 2012年シーズン

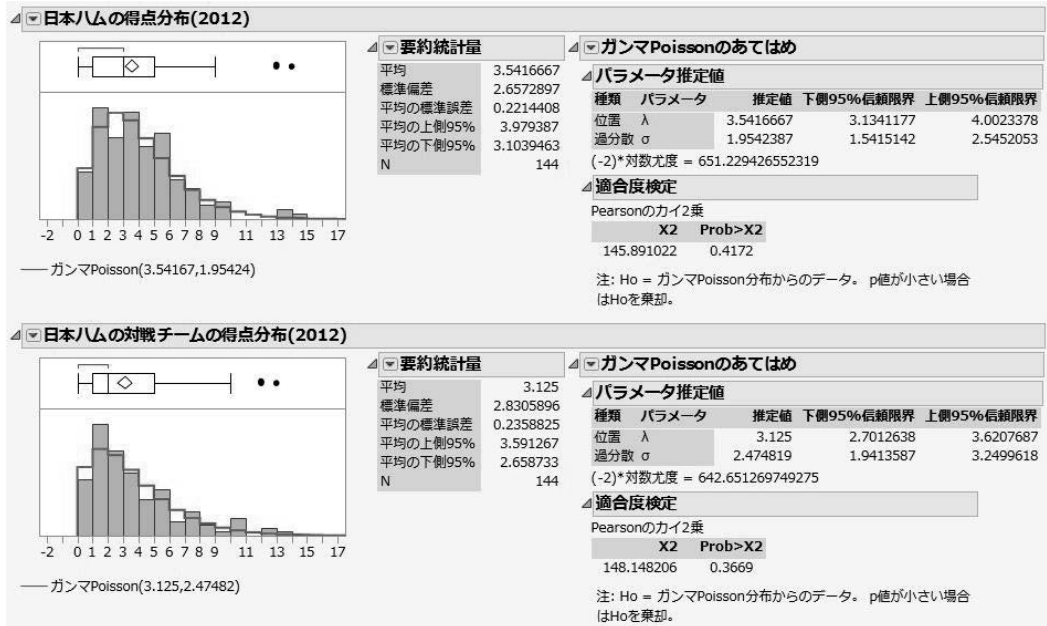


図 10 2位:西武 2012 年シーズン

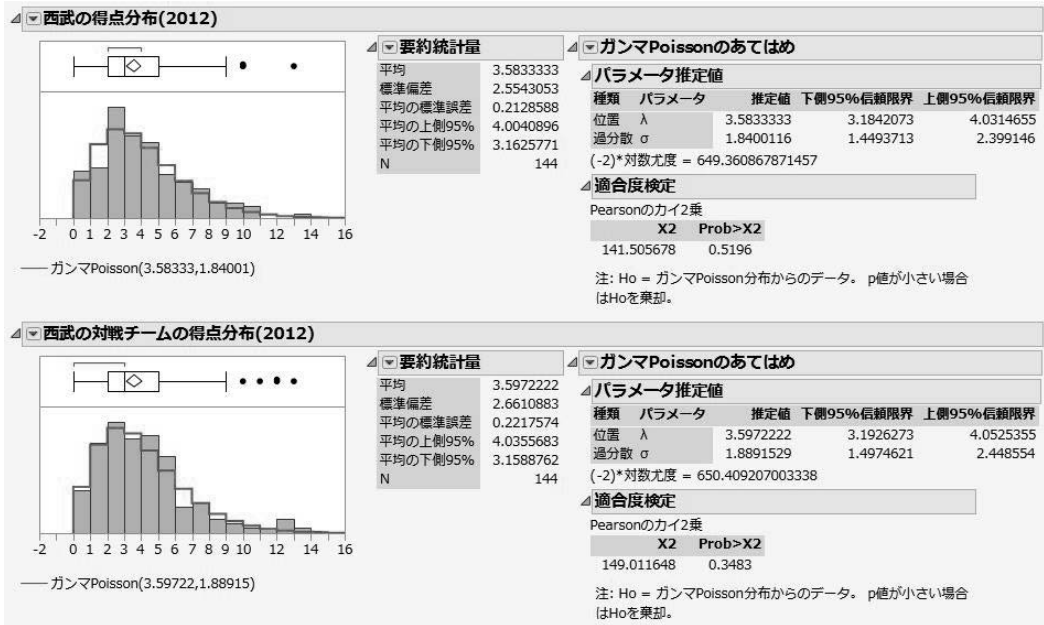


図 11 3位:ソフトバンク 2012 年シーズン

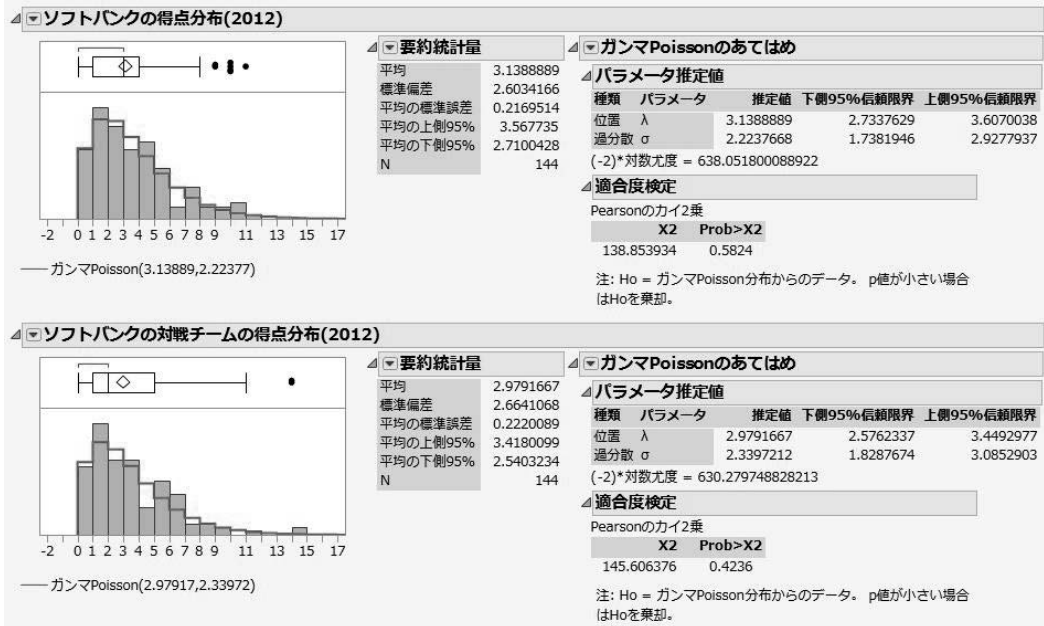


図 12 4位:楽天 2012年シーズン

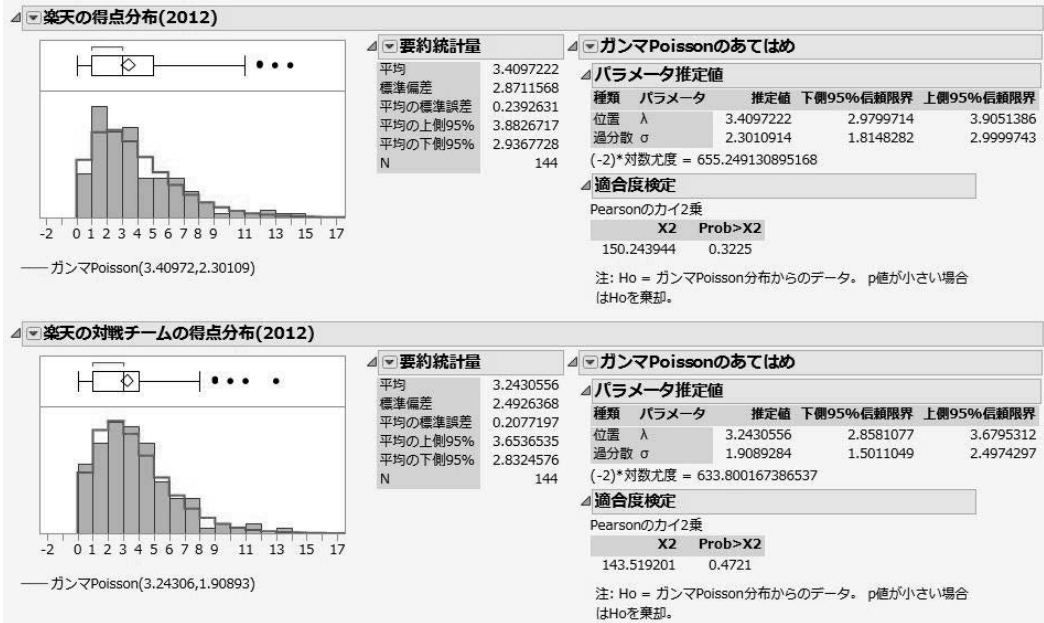


図 13 5位:ロッテ 2012年シーズン

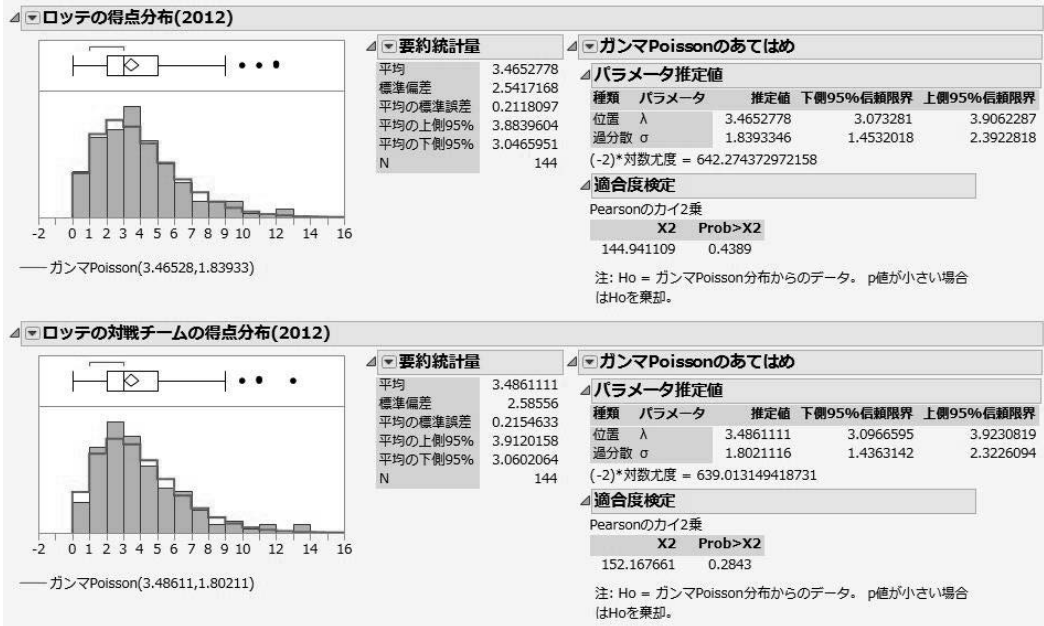
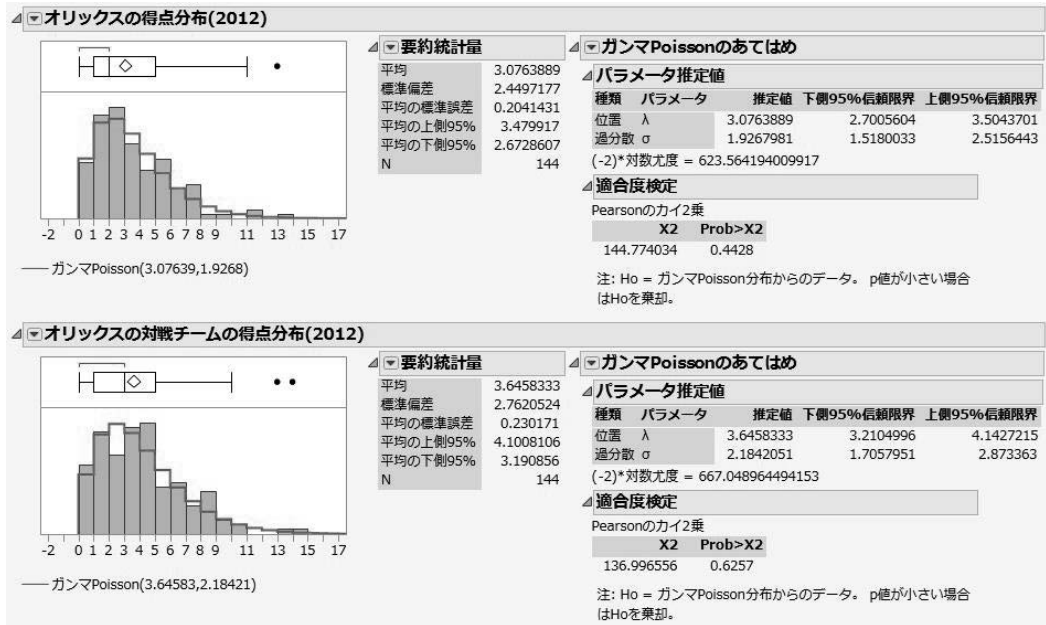


図 14 6位:オリックス 2012年シーズン



4 2013年シーズン

4.1 セリーグ 2013年

図 16 から図 21 を見れば分かるように、6 球団の 1 試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

このシーズンのセリーグの得失点の関係を示す、図 15 を見てすぐに気づくことは、得失点を表す点が右上の方向に移動しているということだ。最近話題になった「飛ぶボール」と「飛ばないボール」の問題があるが、この年のシーズンは「飛ぶボール」が使われたと言っても良いだろう。

このシーズンも昨年 2012 年シーズンに引き続き、投打のバランスの良い巨人の優勝に終わった。興味深いのが DeNA の得点力の高さであり、優勝した巨人を抑えてトップになっている。投手力が強化されれば面白いチームへと成長するだろう。阪神と中日は得点力がほぼ同じだが、投手力が順位を決めたといえる。ヤクルトは相変わらず得点力は高いが、失点が多く最下位に甘んじている。

図 15 セリーグ 2013 年シーズン平均得失点散布図

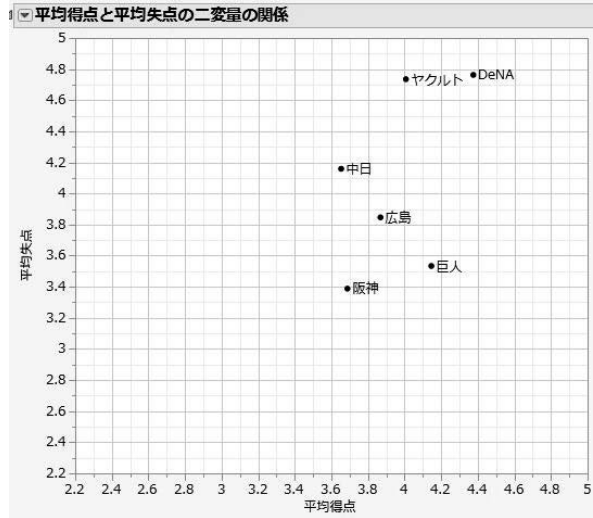


図 16 1位:巨人 2013 年シーズン

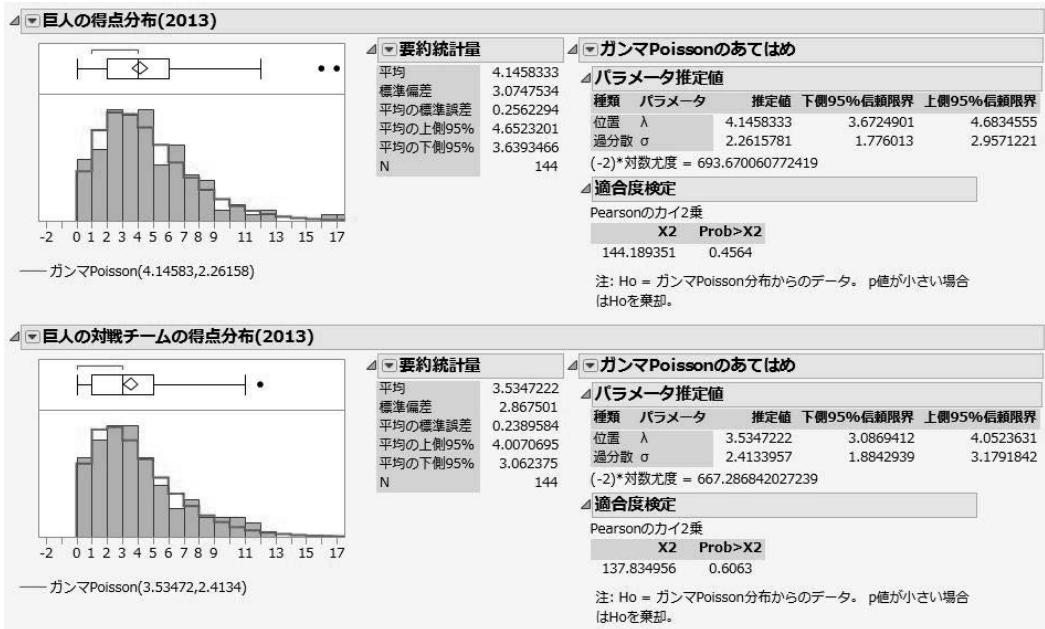


図 17 2位:阪神 2013 年シーズン

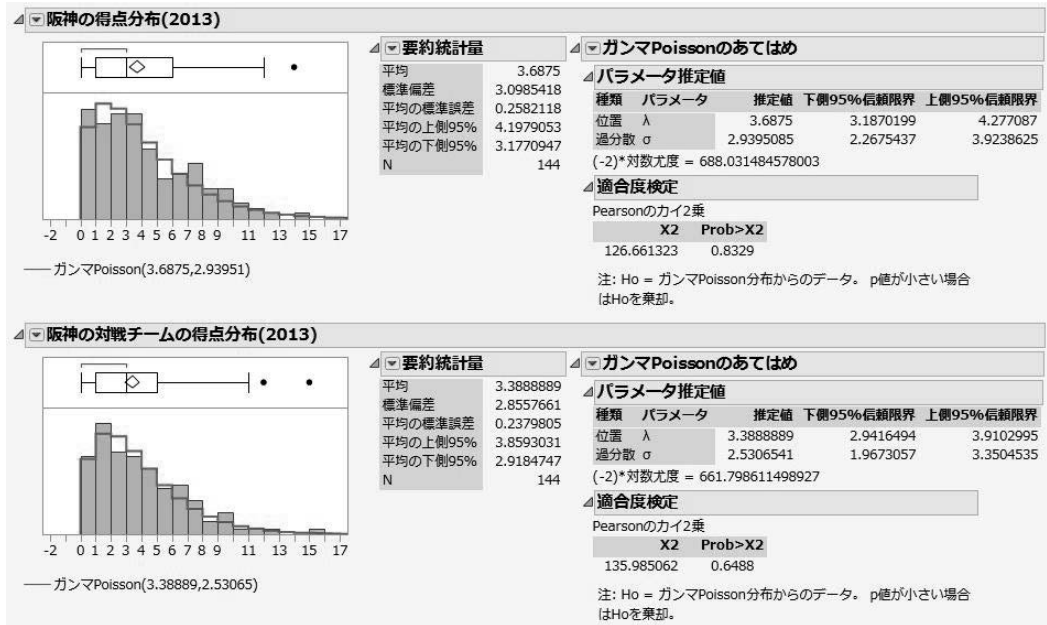


図 18 3位:広島 2013 年シーズン

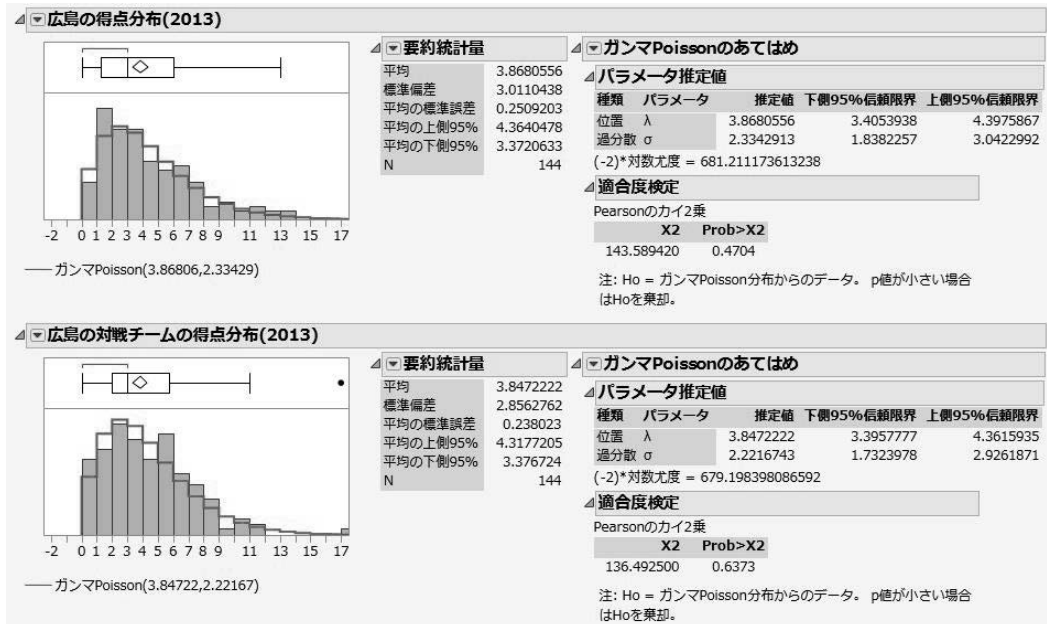


図 19 4位:中日 2013年シーズン

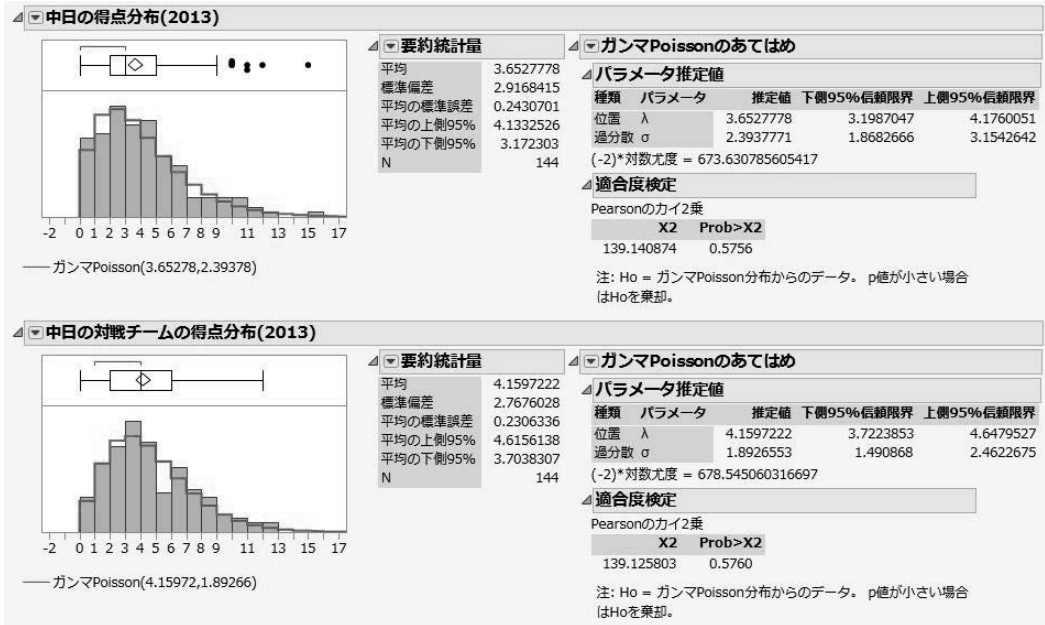


図 20 5位:DeNA 2013年シーズン

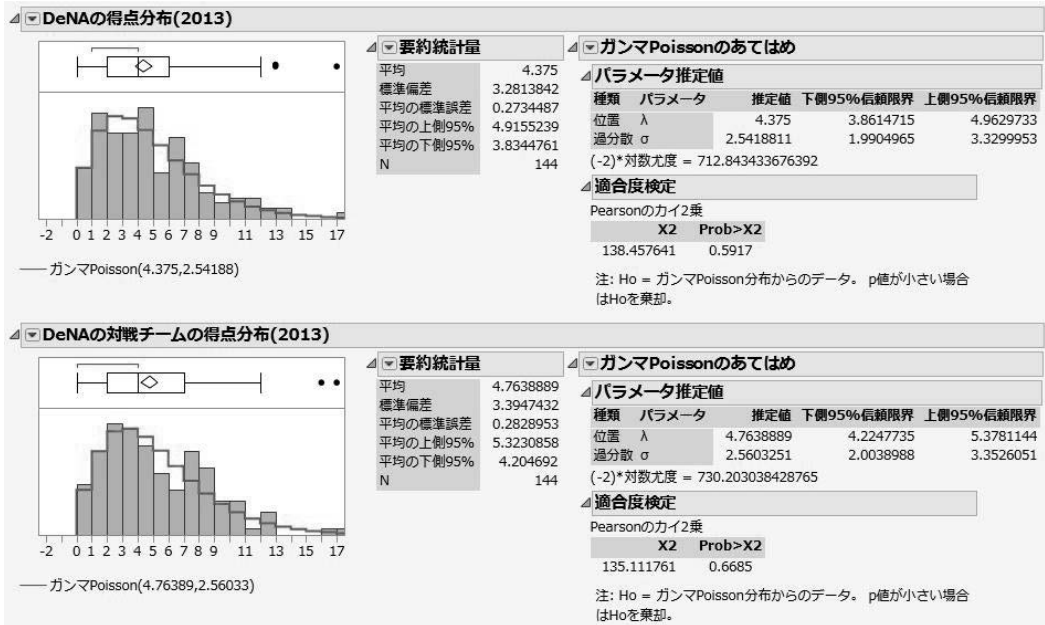
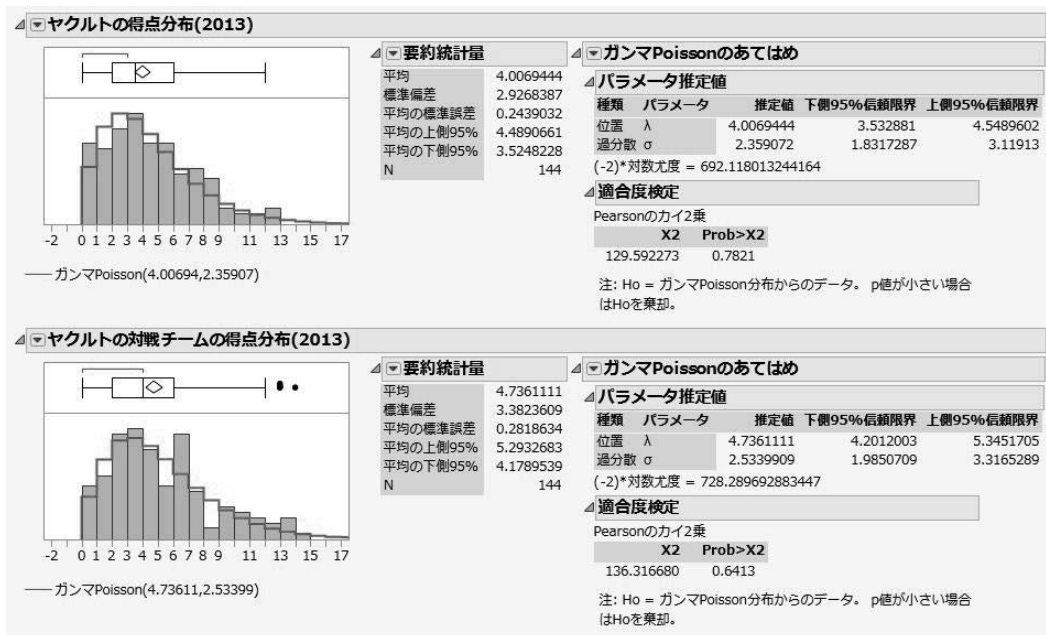


表3 セリーグ・ペナントレース結果 (2013)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①巨人	144	84	53	7	.613	-
②阪神	144	73	67	4	.521	12.5
③広島	144	69	72	3	.489	17.0
④中日	144	64	77	3	.454	22.0
⑤DeNA	144	64	79	1	.448	23.0
⑥ヤクルト	144	57	83	4	.407	28.5

図21 6位:ヤクルト2013年シーズン



4.2 パリーグ 2013 年

図23から図28を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図22を見ると、攻撃力に勝るソフトバンクをかわして、守備力に勝る楽天が優勝しているように見える。しかし、不思議なのはソフトバンクと西武とロッテの順位である。投手力は同等だが打撃力で勝るソフトバンクが西武やロッテよりも下位に位置している。ソフトバンクは試合運びが下手なのかもしれない。この問題については別の機会で議論する。

図 22 パリーグ 2013 年シーズン平均得失点散布図

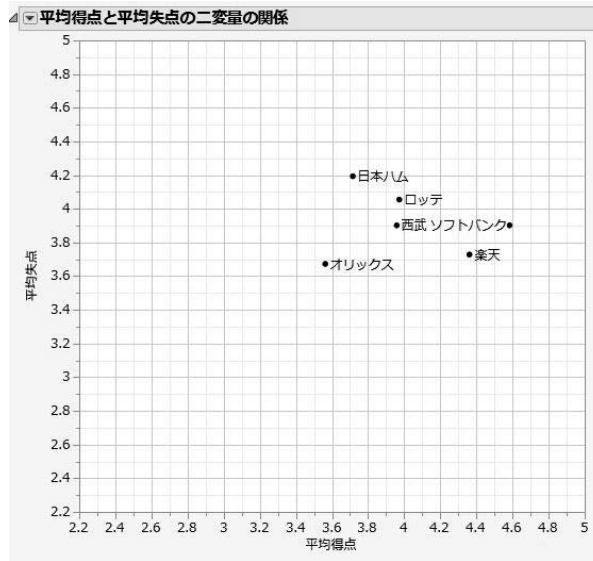


図 23 1 位:楽天 2013 年シーズン

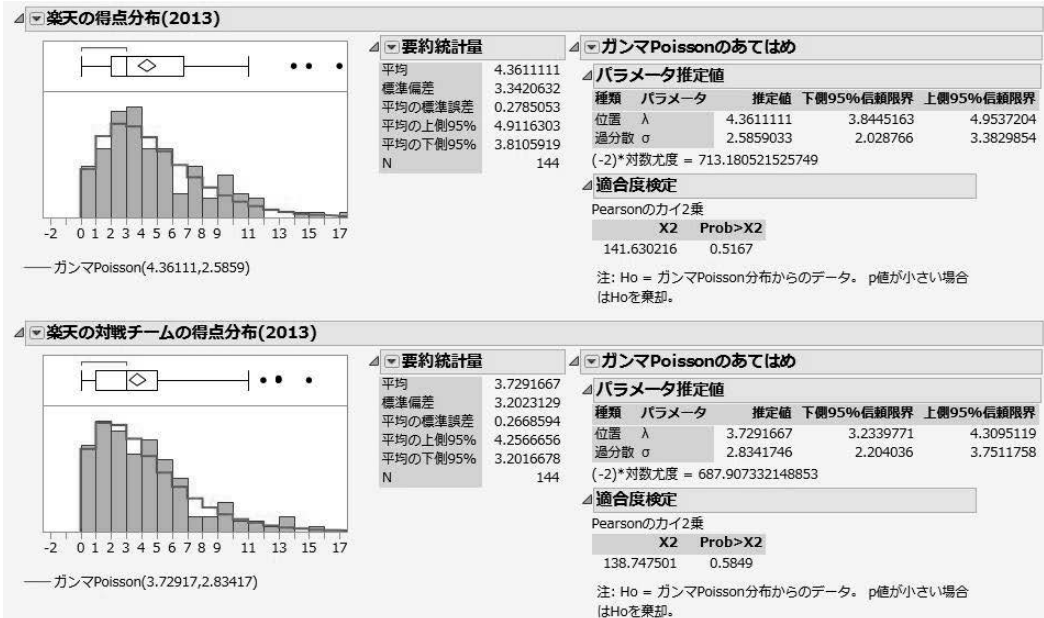


図 24 2位:西武 2013 年シーズン

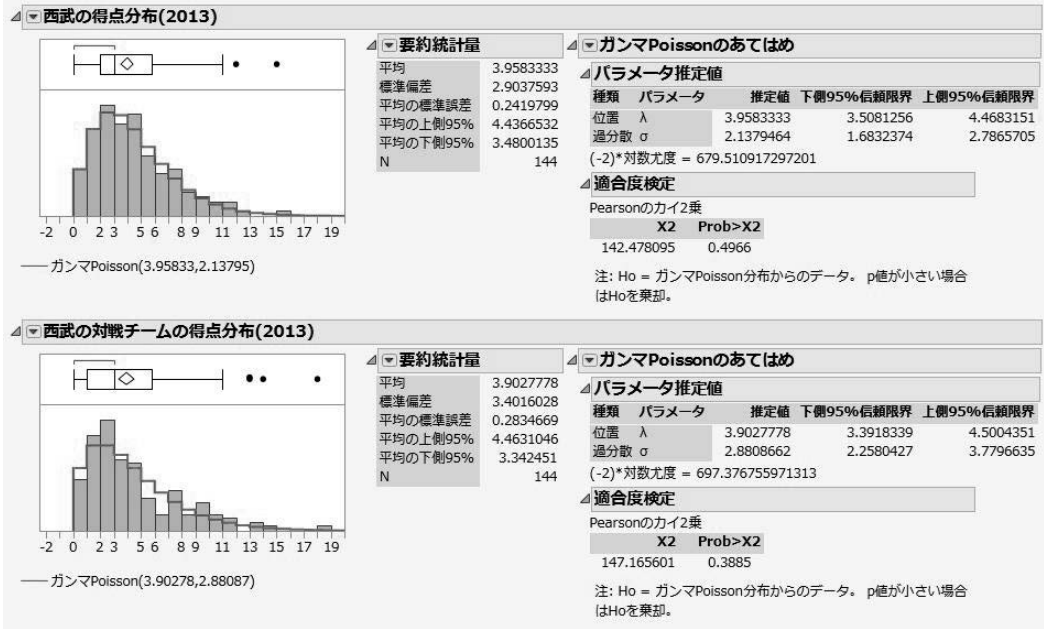


図 25 3位:ロッテ 2013 年シーズン

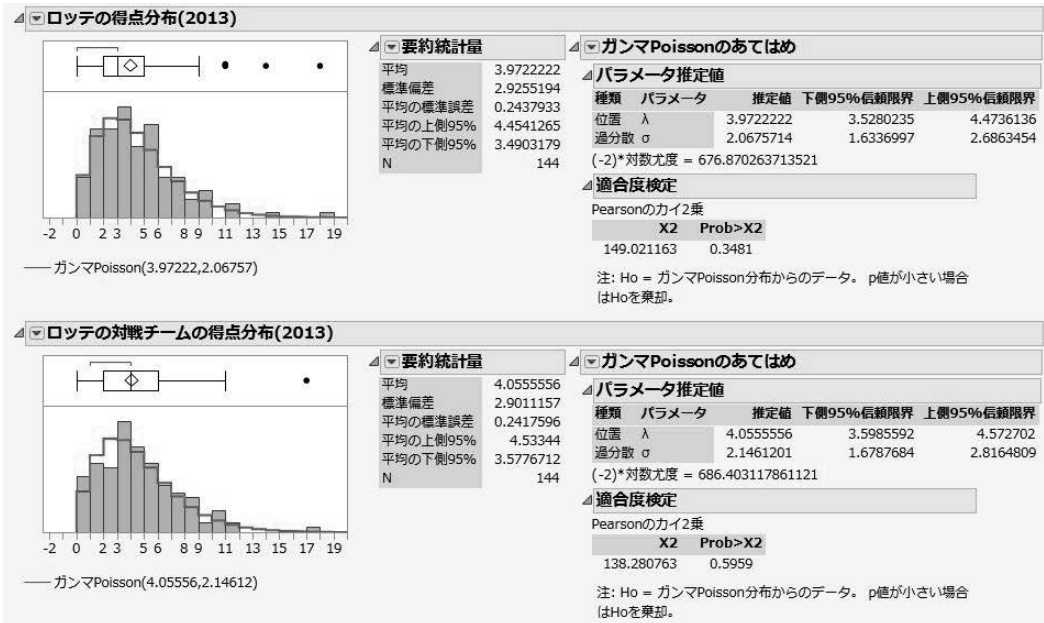


図 26 4位:ソフトバンク 2013年シーズン

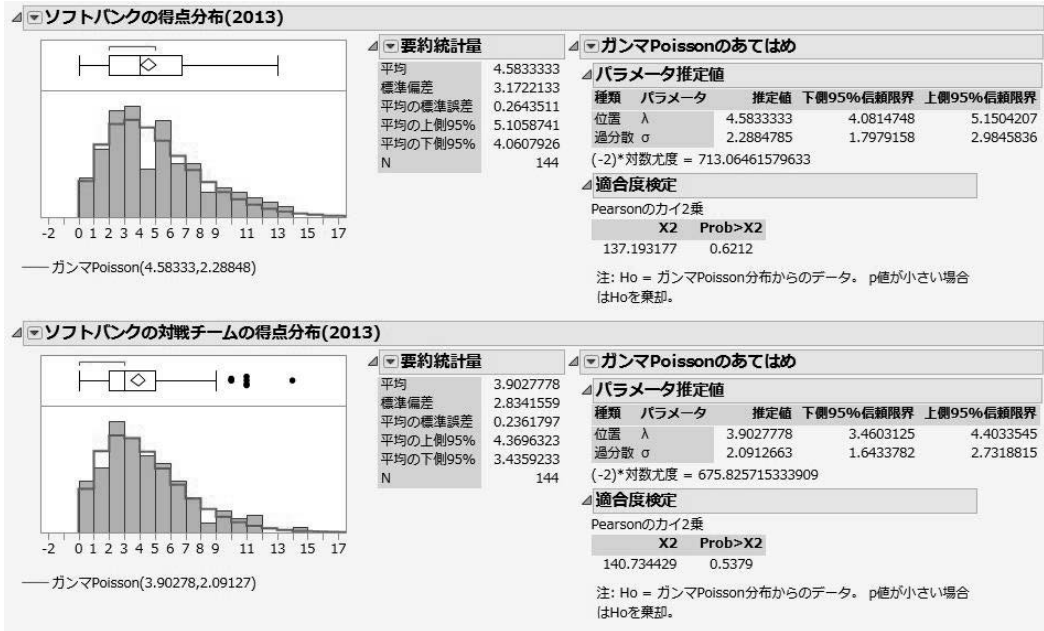
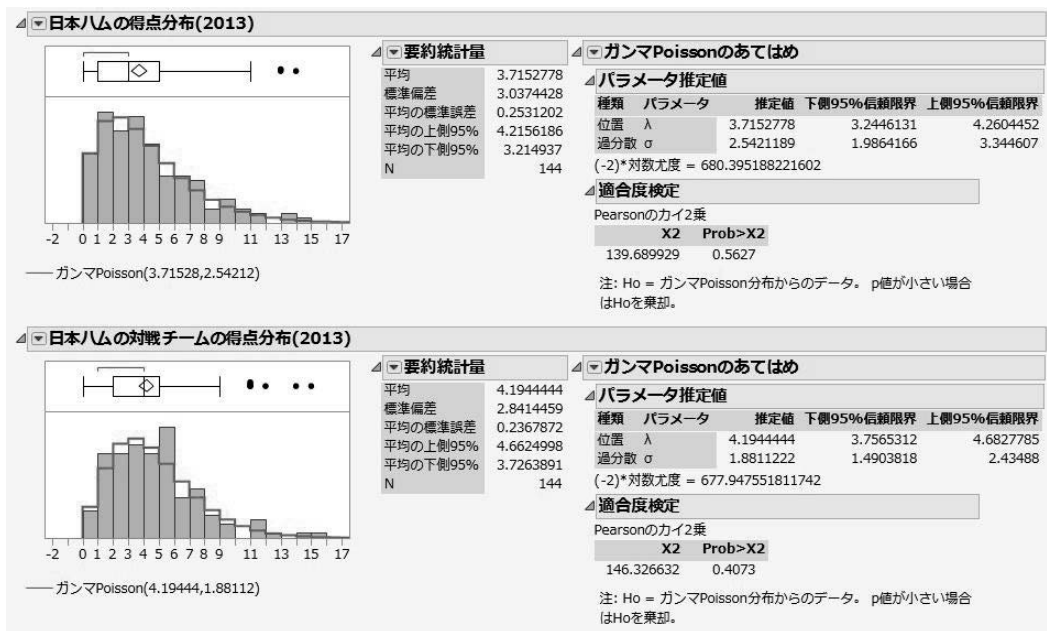


表4 パリーグ・ペナントレース結果 (2013)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①楽天	144	82	59	3	.582	-
②西武	144	74	66	4	.529	7.5
③ロッテ	144	74	68	2	.521	8.5
④ソフトバンク	144	73	69	2	.514	9.5
⑤オリックス	144	66	73	5	.475	15.0
⑥日本ハム	144	64	78	2	.451	18.5

図28 6位:日本ハム 2013年シーズン



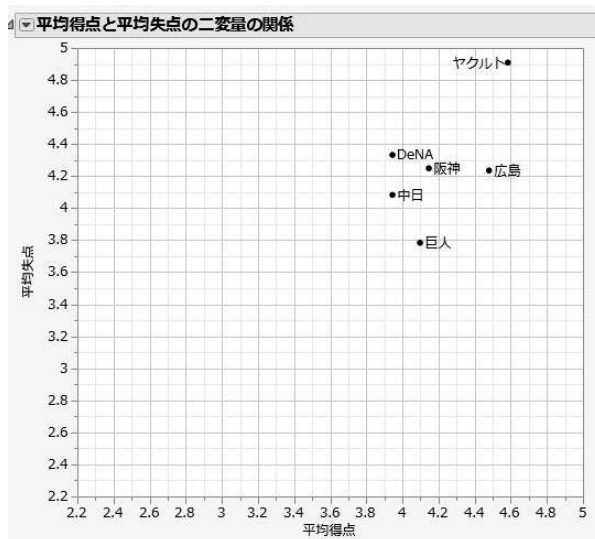
5 2014年シーズン

5.1 セリーグ 2014年

図30から図35を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図29を見れば明らかに、このシーズンも「飛ぶボール」が使われていたと言え、点が右上に集まっている。巨人の優勝は守備力のおかげで、打線は振るわなかった。広島は阪神よりも有利な位

図 29 セリーグ 2014 年シーズン平均得失点散布図



置を占めながら、結果的には阪神に続く3位に終わった。ヤクルトは相変わらず得点力が高いが、その分投手力が際立って弱く最下位に低迷している。

表 5 セリーグ・ペナントレース結果 (2014)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①巨人	144	82	61	1	.573	-
②阪神	144	75	68	1	.524	7.0
③広島	144	74	68	2	.521	7.5
④中日	144	67	73	4	.479	13.5
⑤DeNA	144	67	75	2	.472	14.5
⑥ヤクルト	144	60	81	3	.426	21.0

図 30 1位:巨人 2014 年シーズン

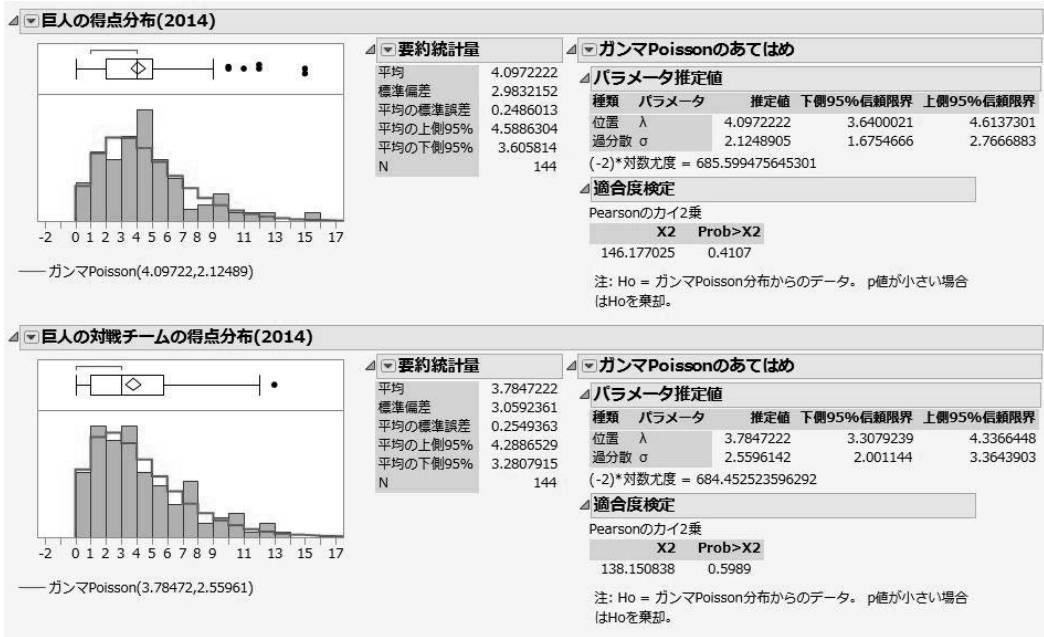


図 31 2位:阪神 2014 年シーズン



図 32 3位:広島 2014年シーズン

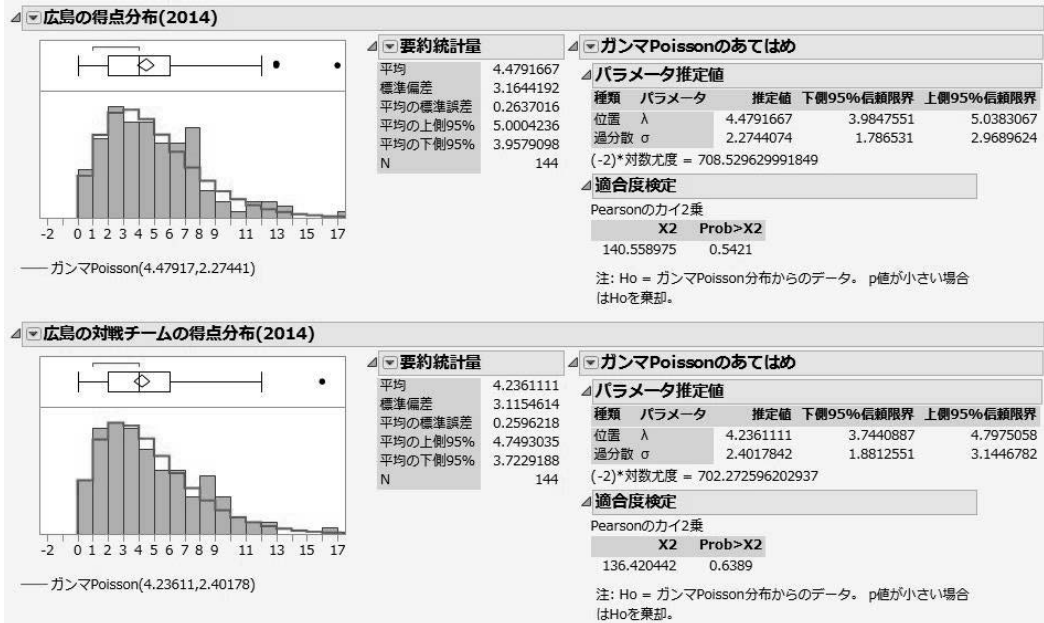


図 33 4位:中日 2014年シーズン

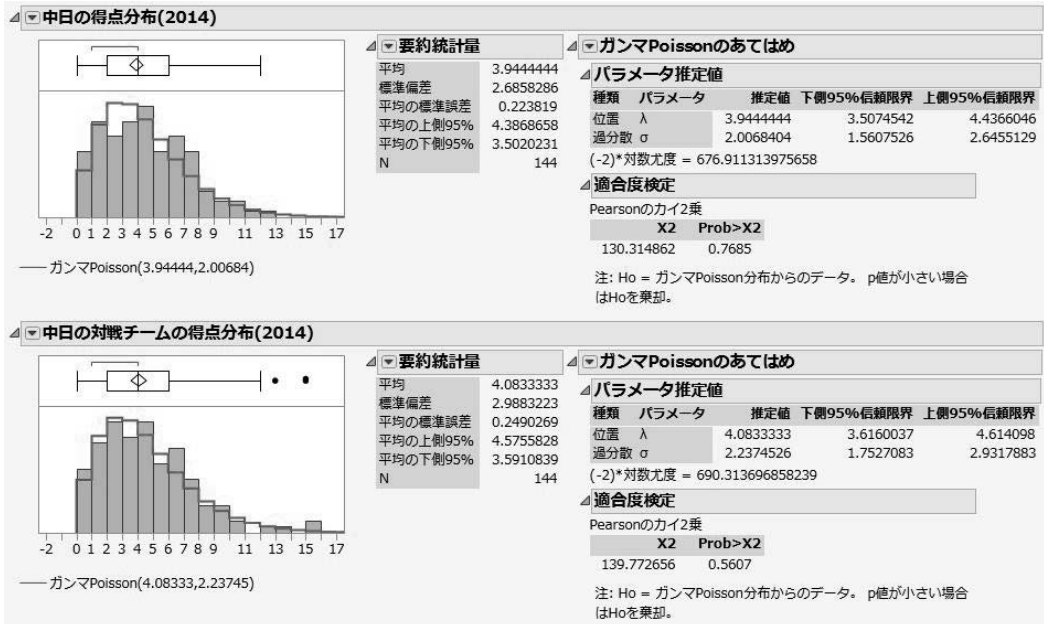


図 34 5位:DeNA2014年シーズン

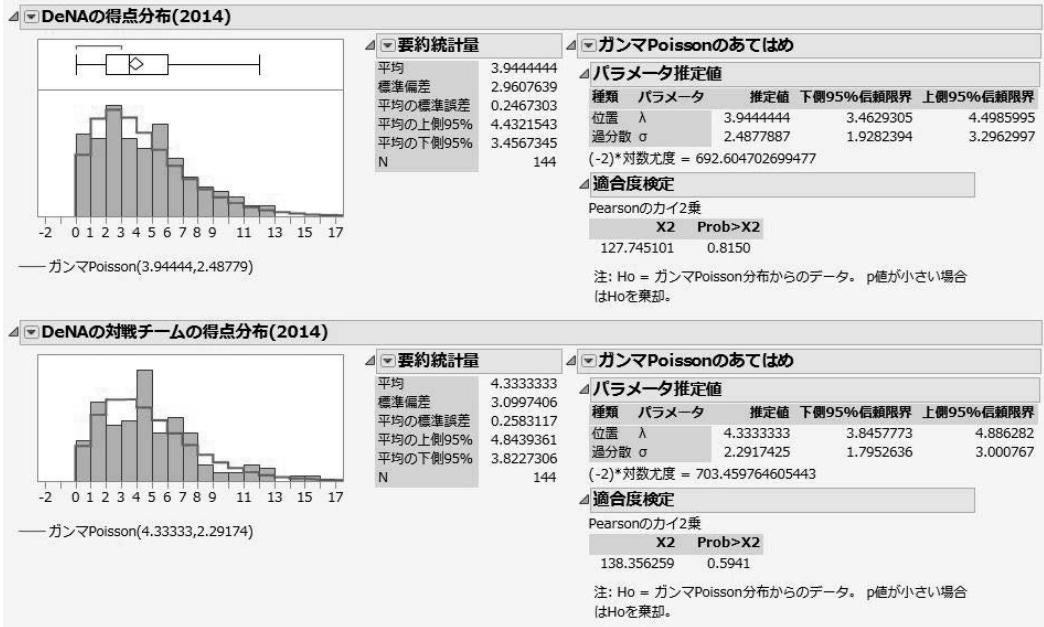


図 35 6位:ヤクルト 2014年シーズン

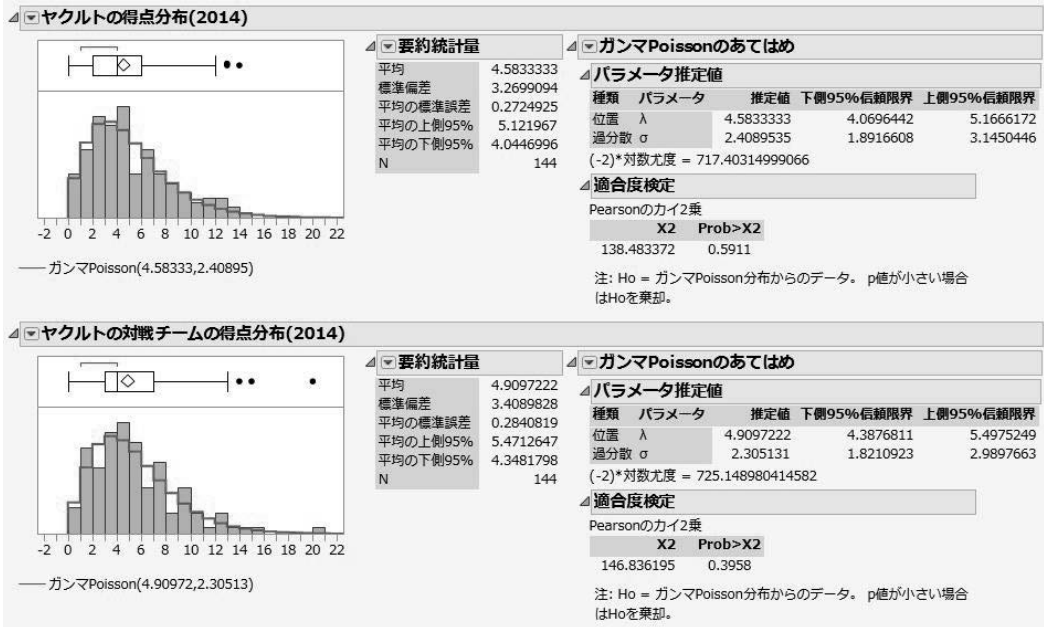
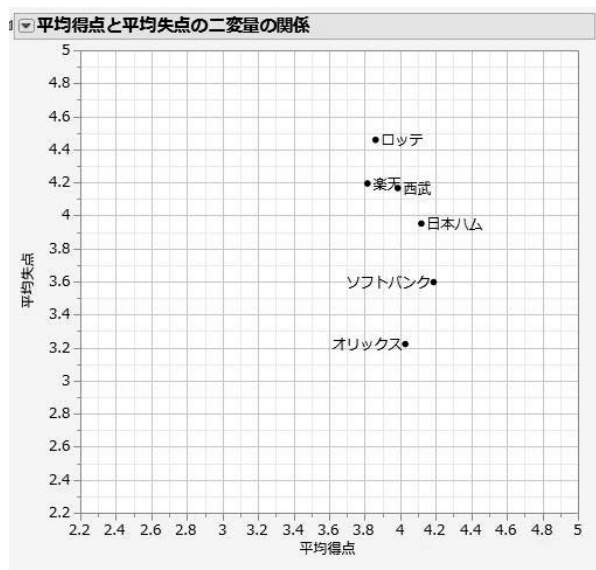


図 36 パリーグ 2014 年シーズン平均得失点散布図



5.2 パリーグ 2014 年

図 37 から図 42 を見れば分かるように、6 球団の 1 試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図 36 を見ると、オリックスが優位に立っているようにみえるが、ゲーム差なしのソフトバンクに勝率で劣り、優勝を逃している。日本ハムの 3 位は順当だが、ロッテが楽天と西武を抑えて 4 位に入っているのが印象に残る。

表 6 パリーグ・ペナントレース結果 (2014)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①ソフトバンク	144	78	60	6	.565	-
②オリックス	144	80	62	2	.563	0.0
③日本ハム	144	73	68	3	.518	6.5
④ロッテ	144	66	76	2	.465	14.0
⑤西武	144	63	77	4	.450	16.0
⑥楽天	144	64	80	0	.444	17.0

図 37 1位:ソフトバンク 2014年シーズン

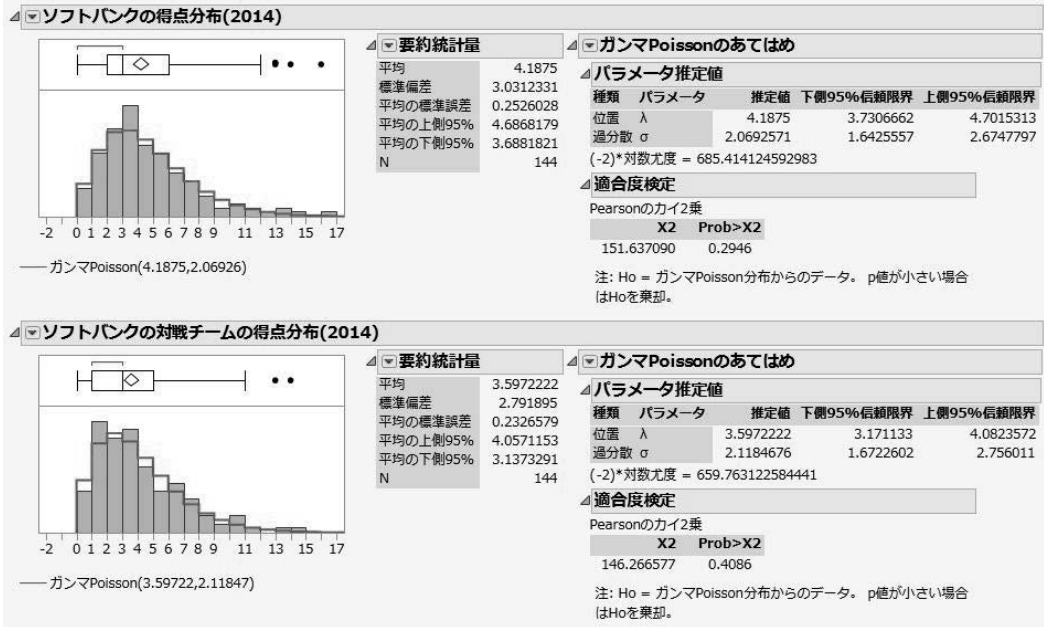


図 38 2位:オリックス 2014年シーズン

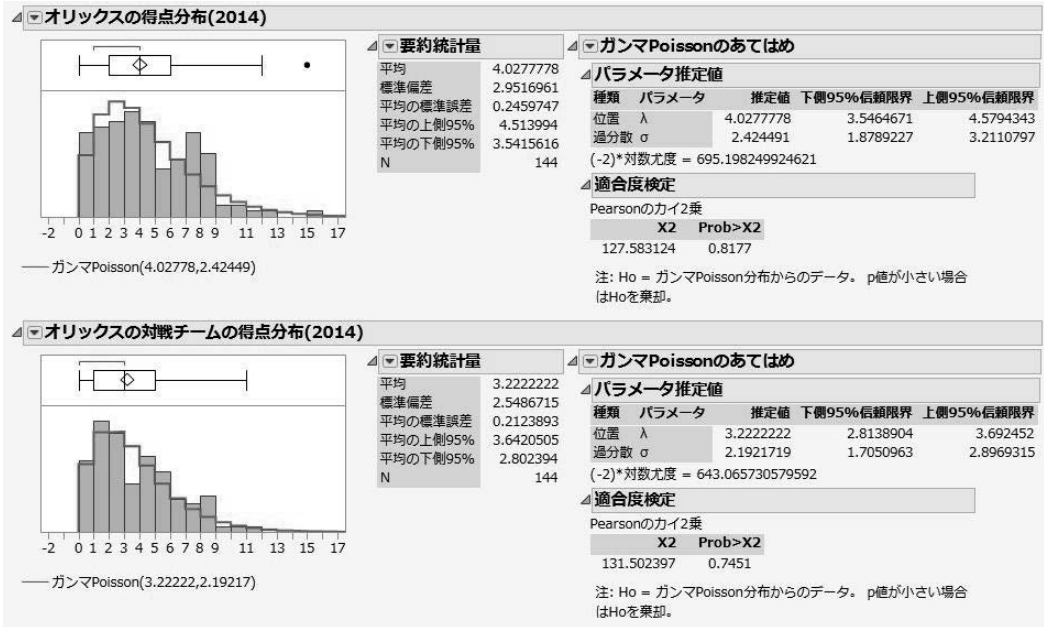


図39 3位:日本ハム 2014年シーズン

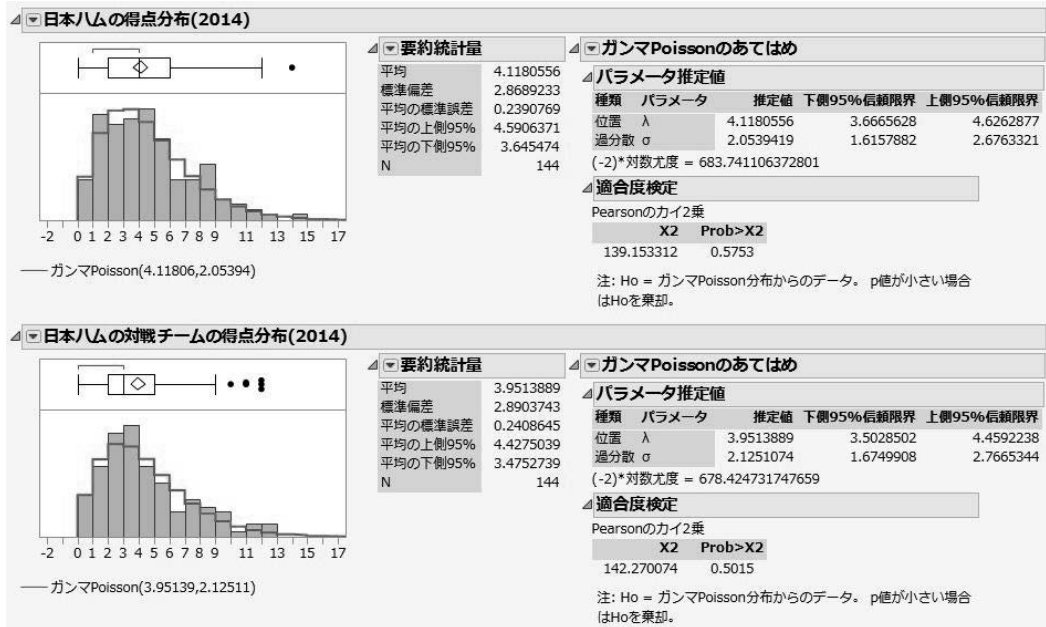


図40 4位:ロッテ 2014年シーズン

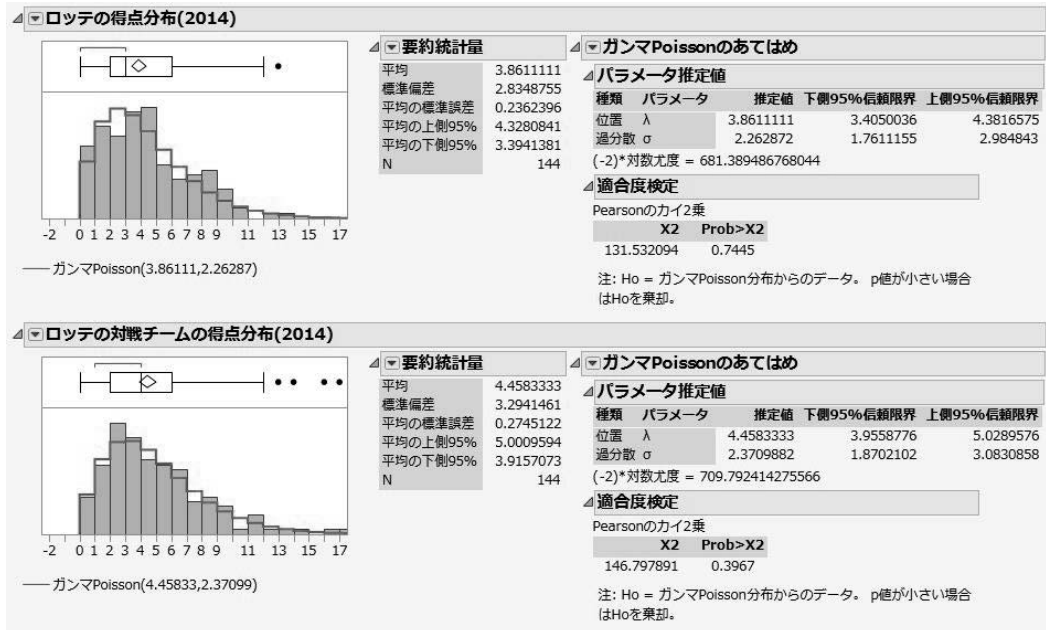
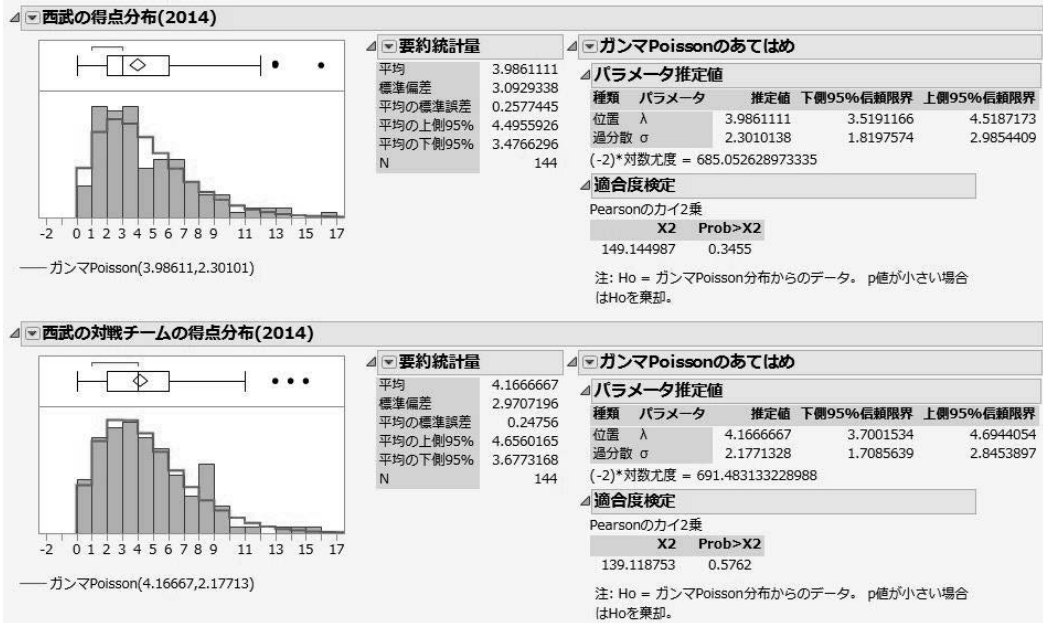


図 41 5位:西武 2014 年シーズン



6 2015年シーズン

6.1 セリーグ2015年

図44から図49を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

このシーズンでは、ヤクルトが打撃が例年より好調なうえに、投手力がますますの成績であることから、優勝しているようだ。2位の巨人は得失点共にすくなく、こじんまりした印象を受ける。阪神は、得点力が最低で、守備力も下から2番目であるにもかかわらず3位につけている。4位の広島は、もっと高い順位になるべき得失点ではあるがその理由は分からない。

ボールも「飛ぶボール」と「飛ばないボール」の観点からみると、6つの点が真ん中を中心にバラついているのが分かる。詳しく調べなければならぬが、「飛ぶボール」を使った試合と「飛ばないボール」を使った試合の両方が混じっているのかもしれない。

表7 セリーグ・ペナントレース結果(2015)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①ヤクルト	143	76	65	2	.539	-
②巨人	143	75	67	1	.528	1.5
③阪神	143	70	71	2	.496	6.0
④広島	143	69	71	3	.493	6.5
⑤中日	143	62	77	4	.446	13.0
⑥DeNA	143	62	80	1	.437	14.5

図 43 セリーグ 2015 年シーズン平均得失点散布図

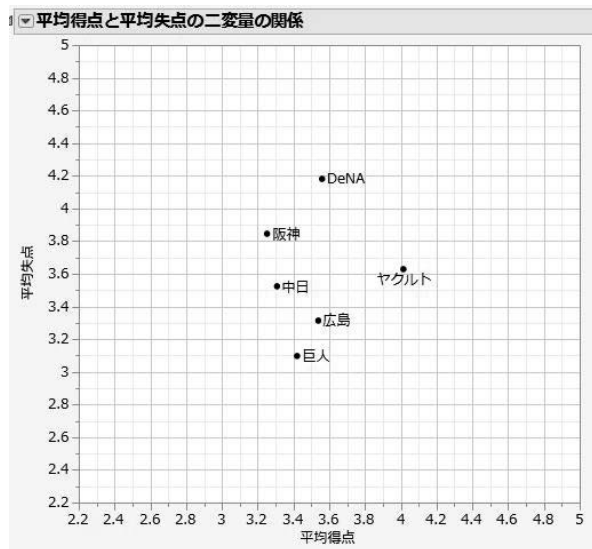


図 44 1 位:ヤクルト 2015 年シーズン

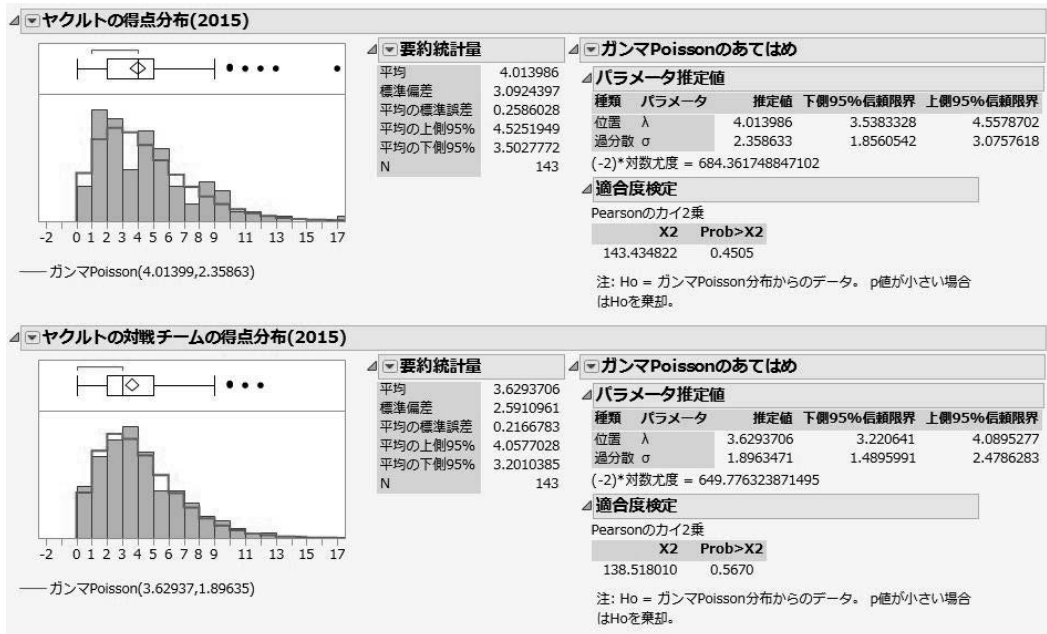


図 45 2位:巨人 2015年シーズン

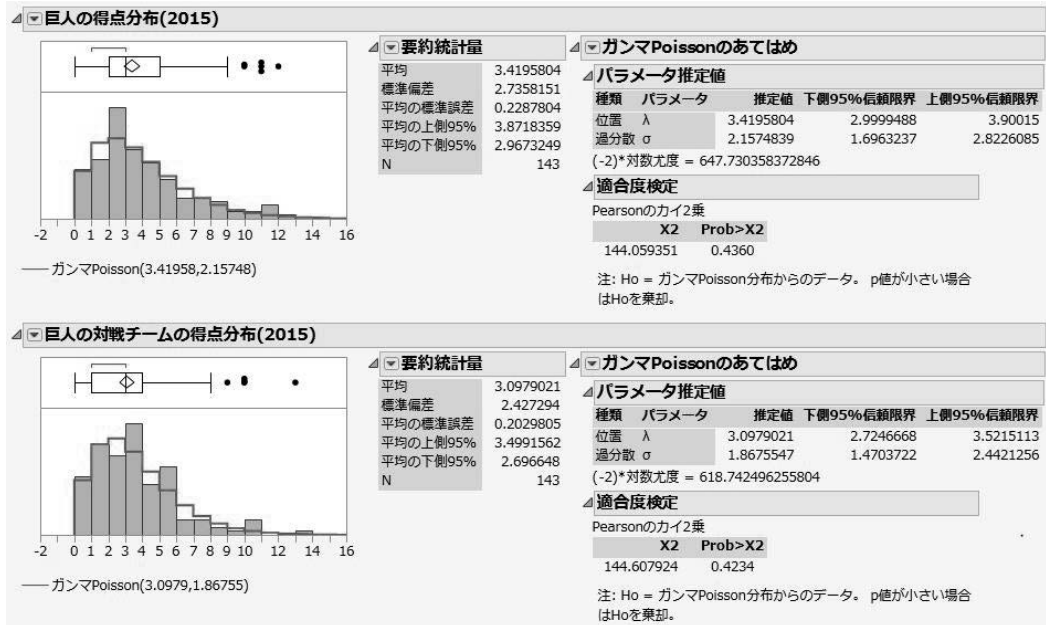


図 46 3位:阪神 2015年シーズン

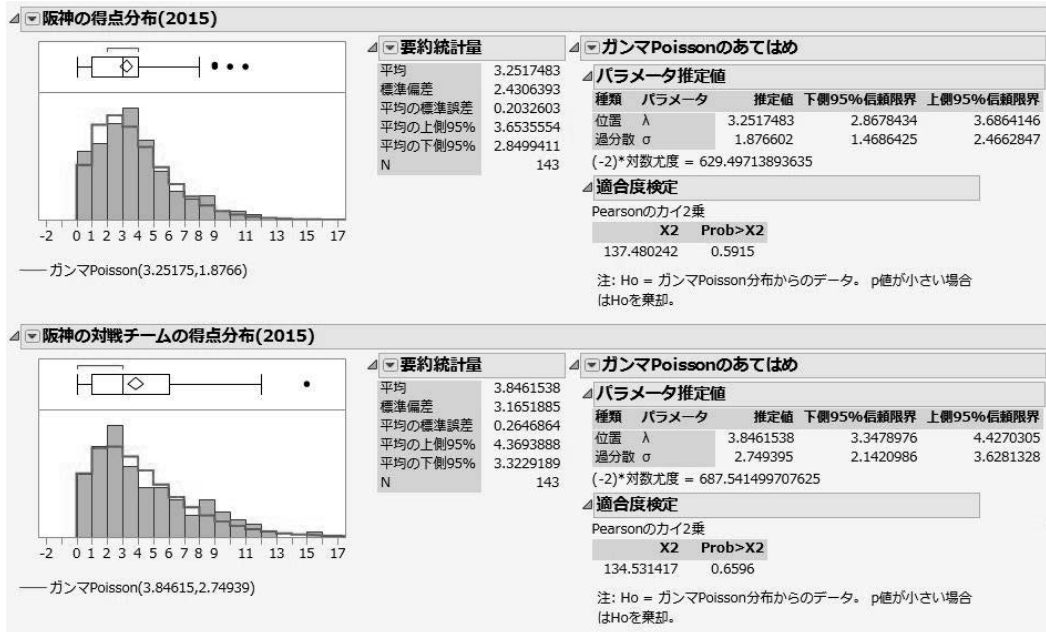


図 47 4位:広島 2015 年シーズン

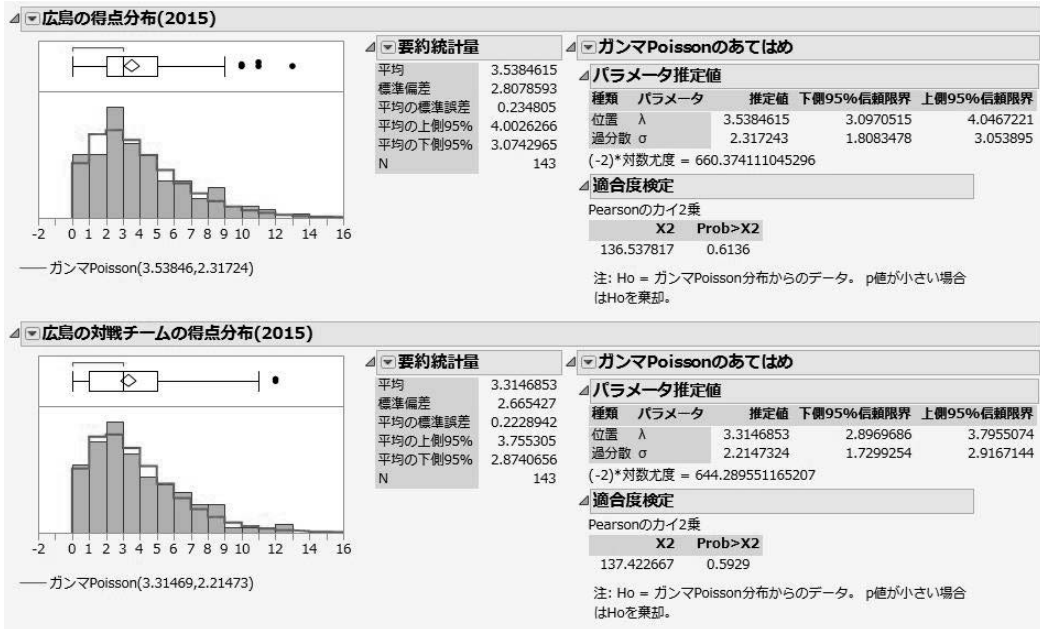


図 48 5位:中日 2015 年シーズン

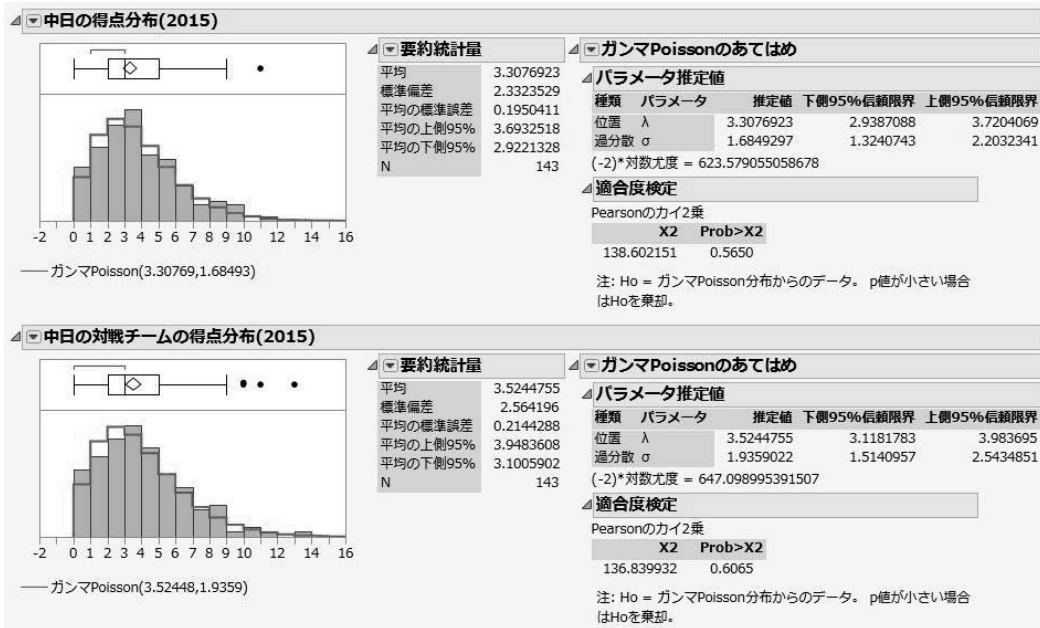
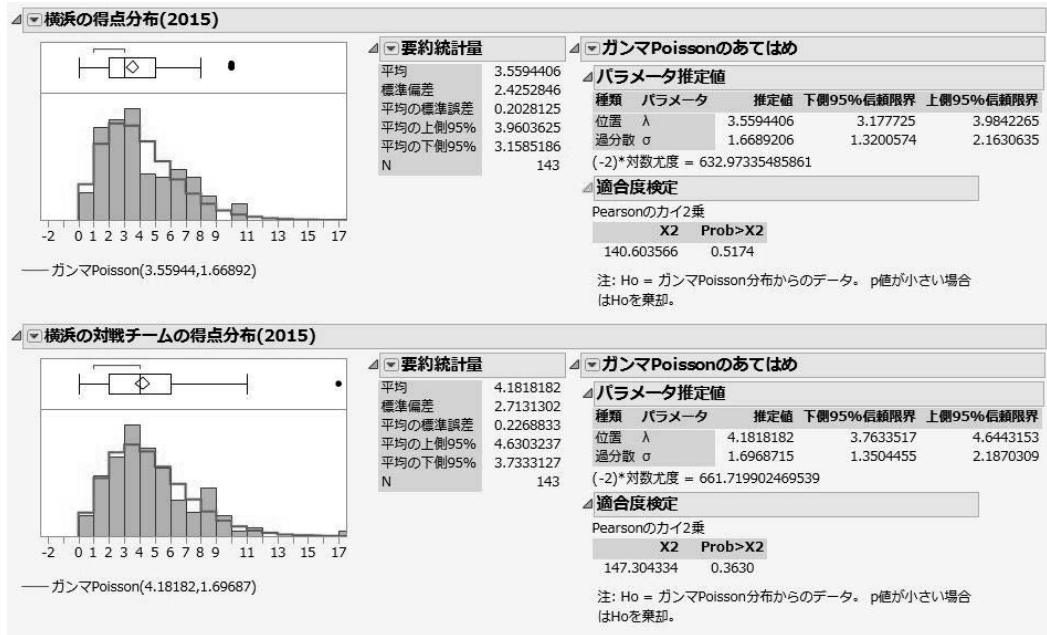


図 49 6位:DeNA2015年シーズン



6.2 パリーグ 2015年

図 51 から図 56 を見れば分かるように、6 球団の 1 試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図 50 を見ればすぐ分かるように、ソフトバンクの圧勝と楽天の惨敗である。図を見る限り西武は日本ハムやオリックスの上位にきてもおかしくないのだが、4 位に終わっている。

表 8 パリーグ・ベナントレース結果 (2015)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①ソフトバンク	143	90	49	4	.647	-
②日本ハム	143	79	62	2	.560	12.0
③ロッテ	143	73	69	1	.514	18.5
④西武	143	69	69	5	.500	20.5
⑤オリックス	143	61	80	2	.433	30.0
⑥楽天	143	57	83	3	.407	33.5

図 50 バリリーグ 2015 年シーズン平均得失点散布図

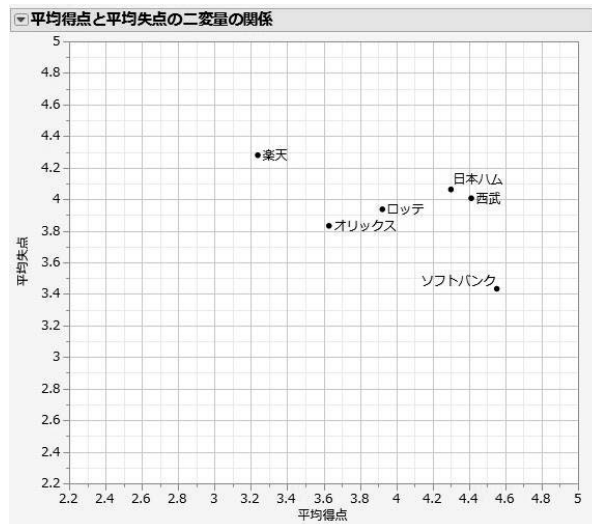


図 51 1 位:ソフトバンク 2015 年シーズン

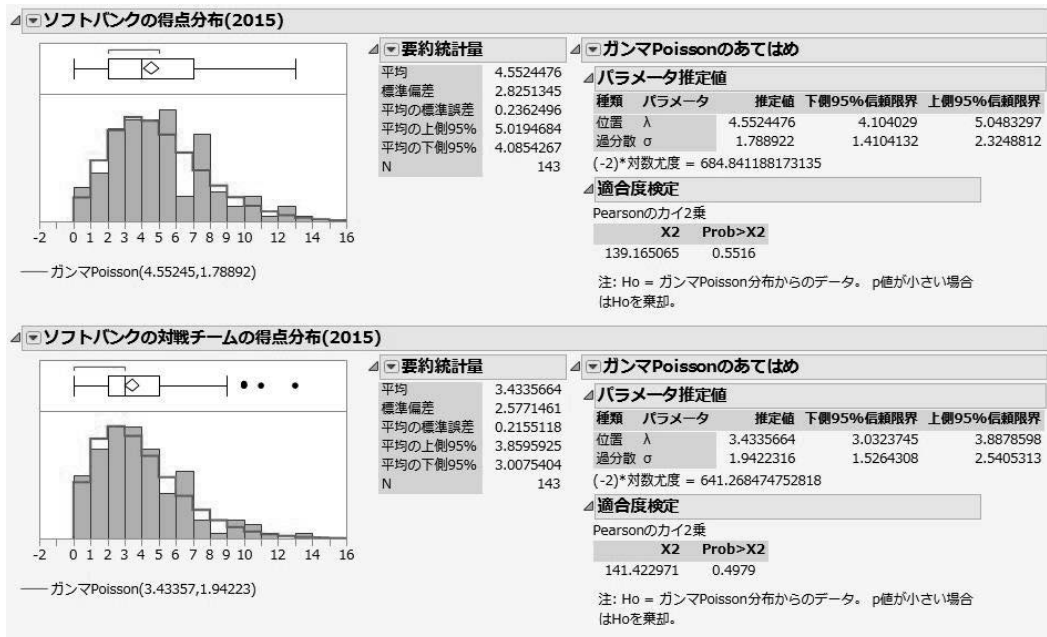


図 52 2位:日本ハム 2015年シーズン

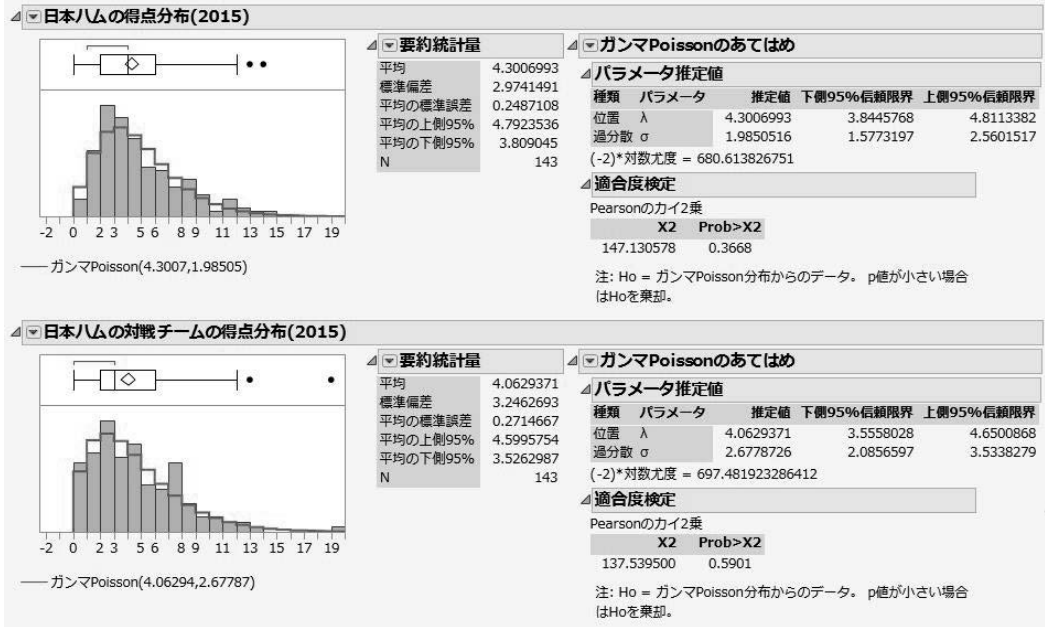


図 53 3位:ロッテ 2015年シーズン

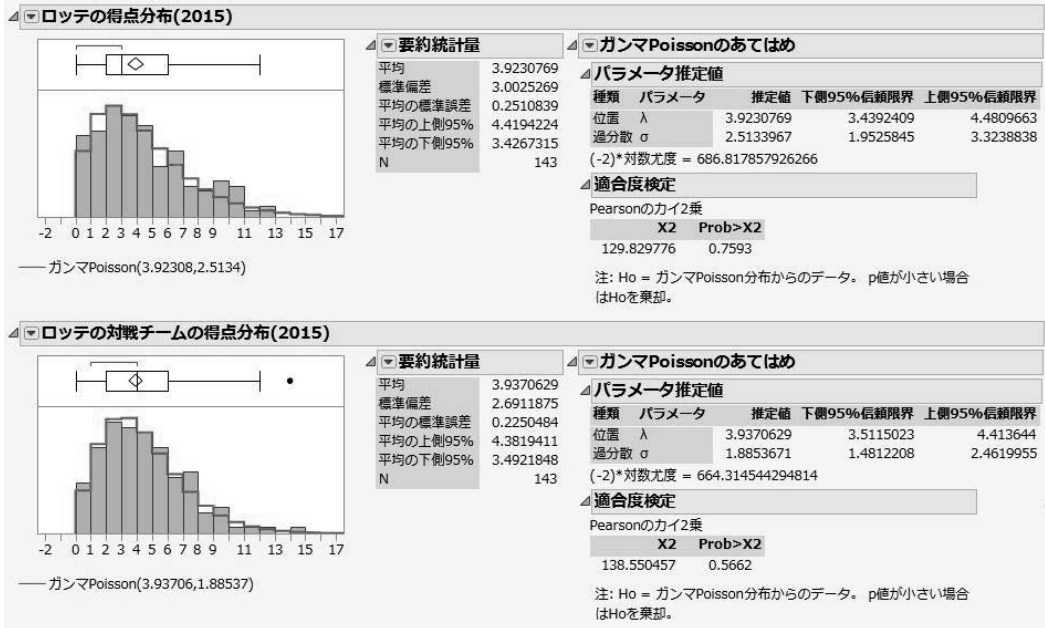


図 54 4位:西武 2015 年シーズン

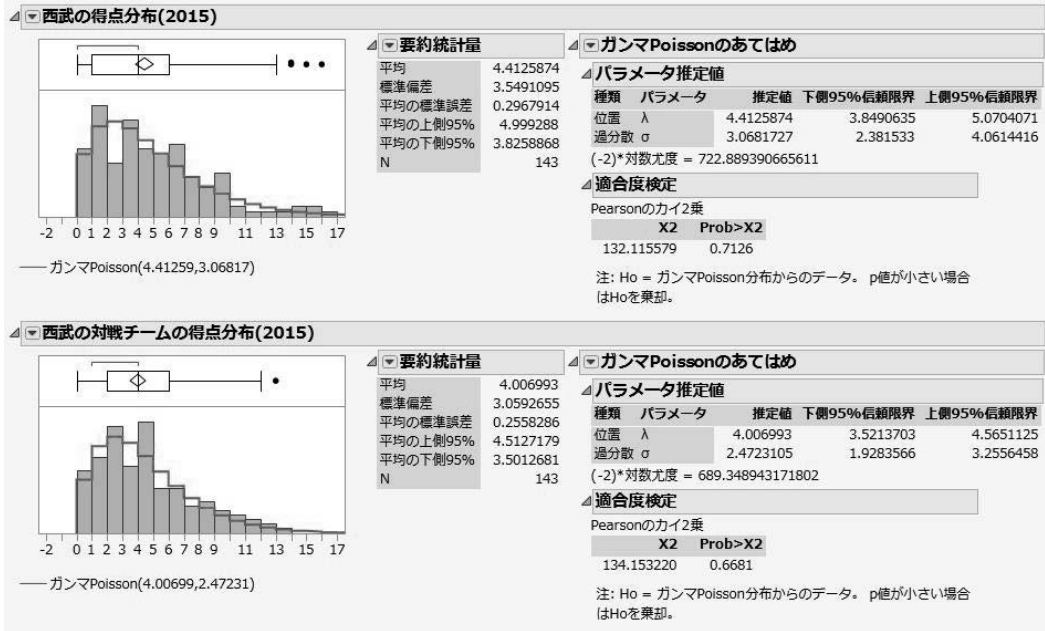


図 55 5位:オリックス 2015 年シーズン

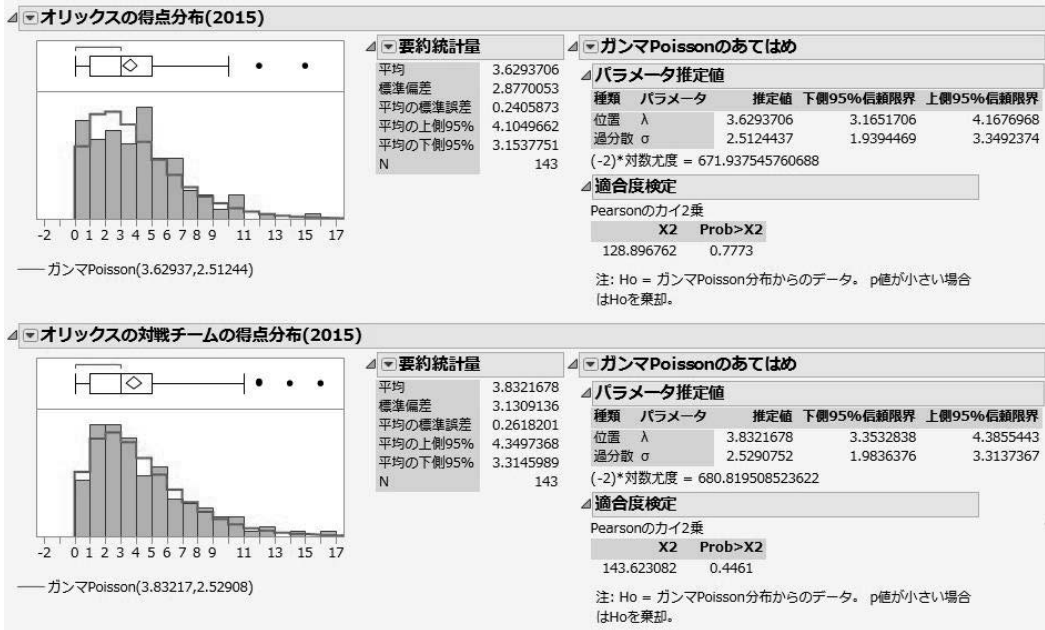
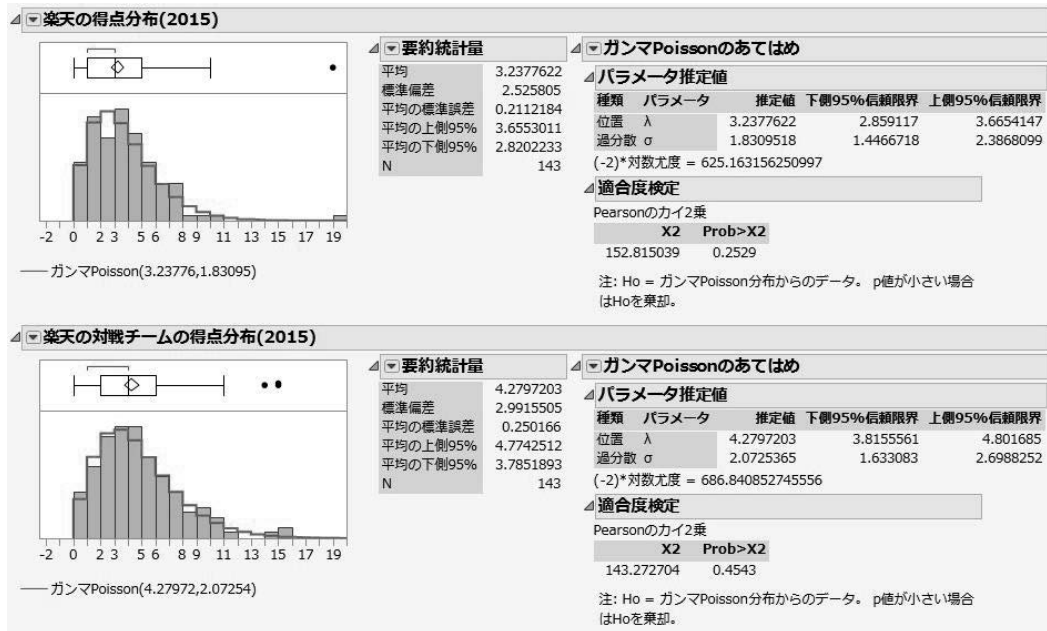


図 56 6位:楽天2015年シーズン



7 2016年シーズン

7.1 セリーグ2016年シーズン

図 58 から図 63 を見れば分かるように、6 球団の 1 試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図 57 から読み取れるように、広島は優勝である。得点力、守備力いずれも 1 位である。2015 年シーズン優勝のヤクルトが 2016 年シーズンに 5 位に甘んじているように、2017 年シーズンの広島は優勝は難しいだろう。このシーズンは、いろいろな要素と運とが重なって優勝したのだろうと推察される。巨人、中日、阪神を示す点は真ん中に集まっている。注目すべきは DeNA だろう。強力な打撃に、守備力の高まりで来シーズンが楽しみなチームである。

図 57 セリーグ 2016 年シーズン平均得失点散布図

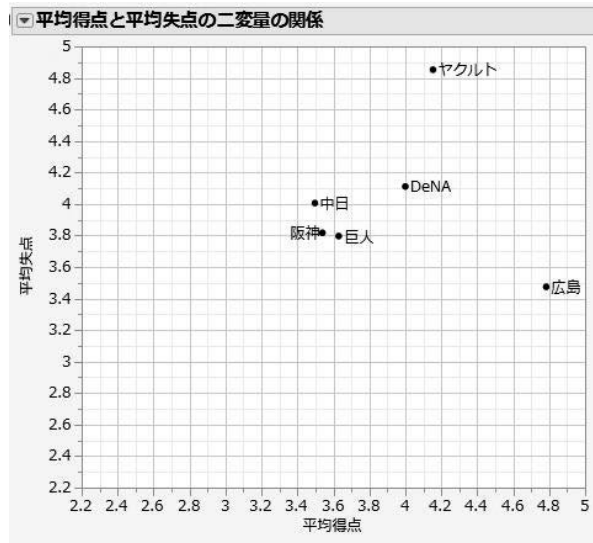


図 58 1 位:広島 2016 年シーズン

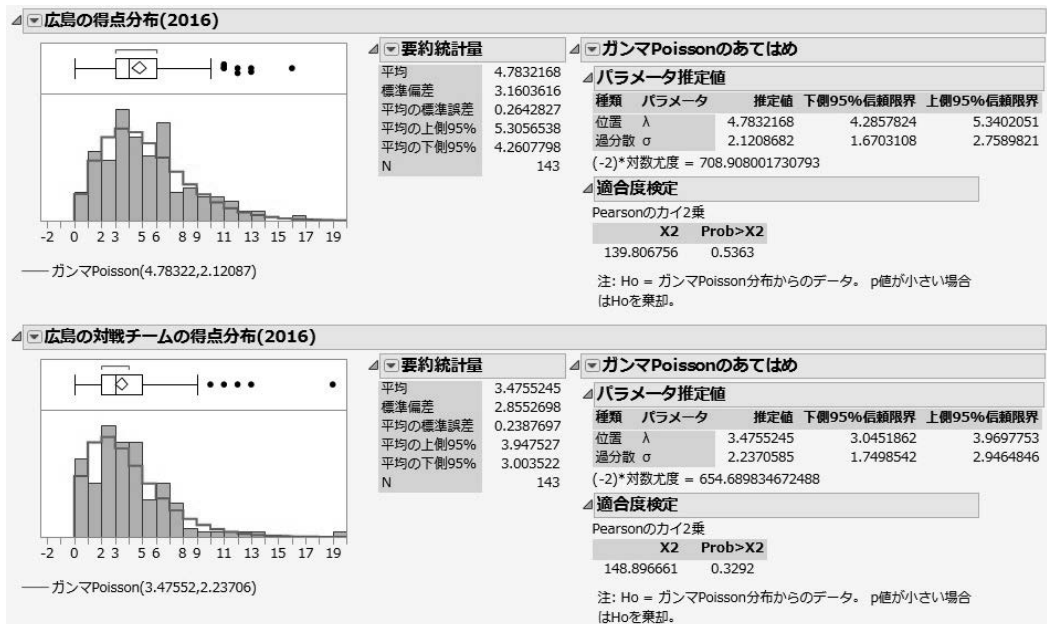


図 59 2位:巨人 2016年シーズン

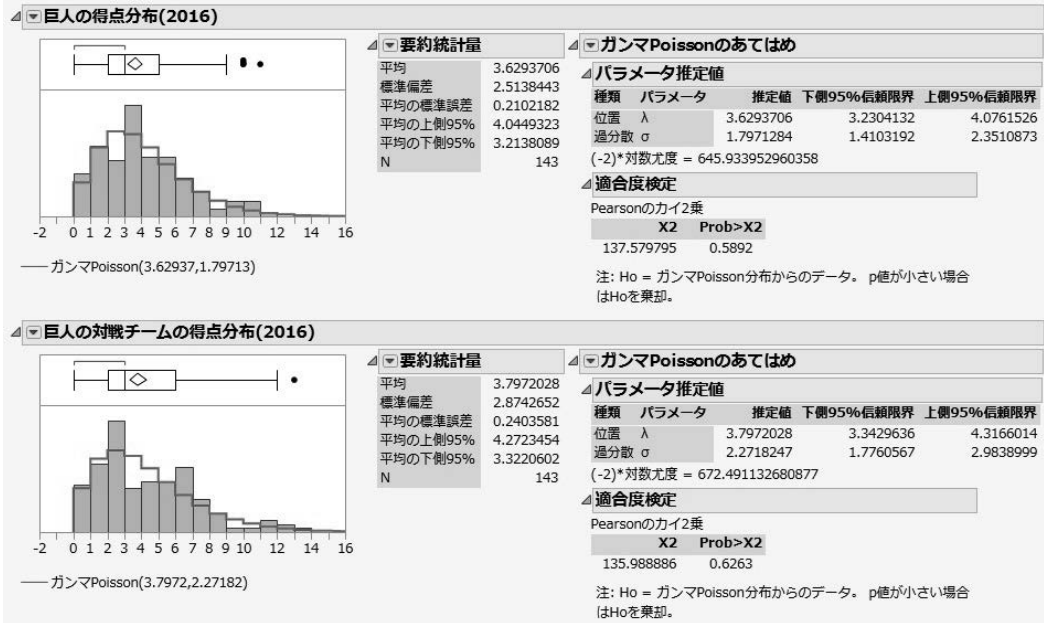


図 60 3位:DeNA 2016年シーズン

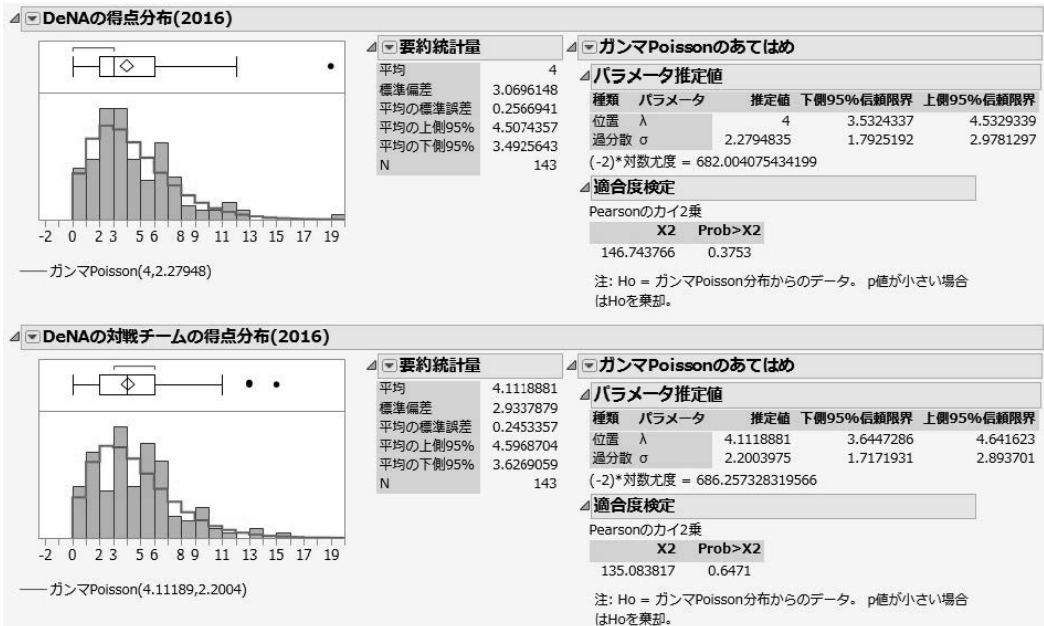


図 61 4位:阪神 2016年シーズン

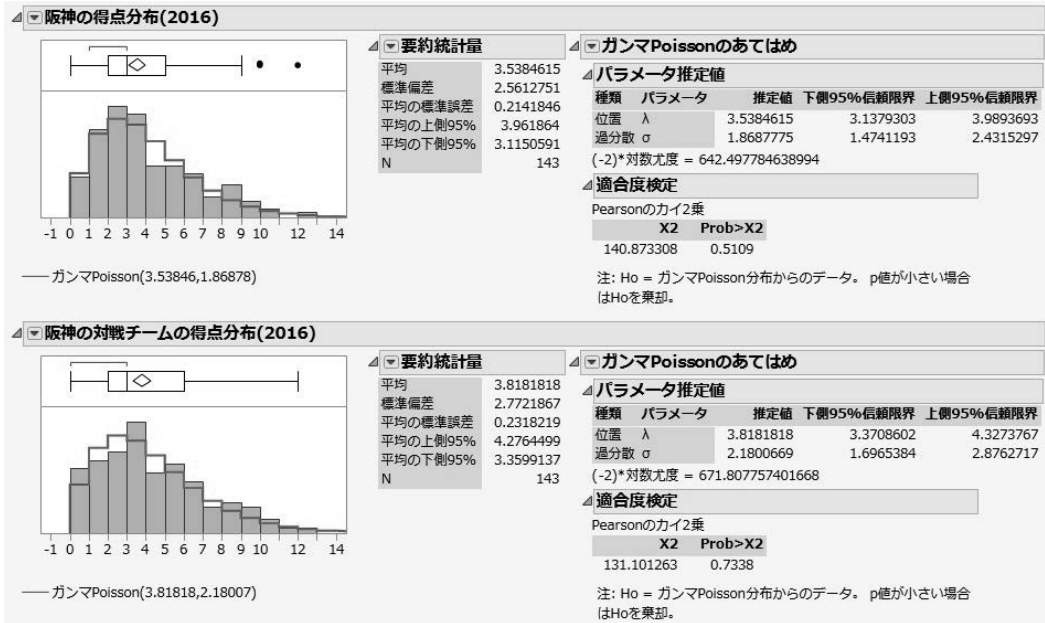


図 62 5位:ヤクルト 2016年シーズン

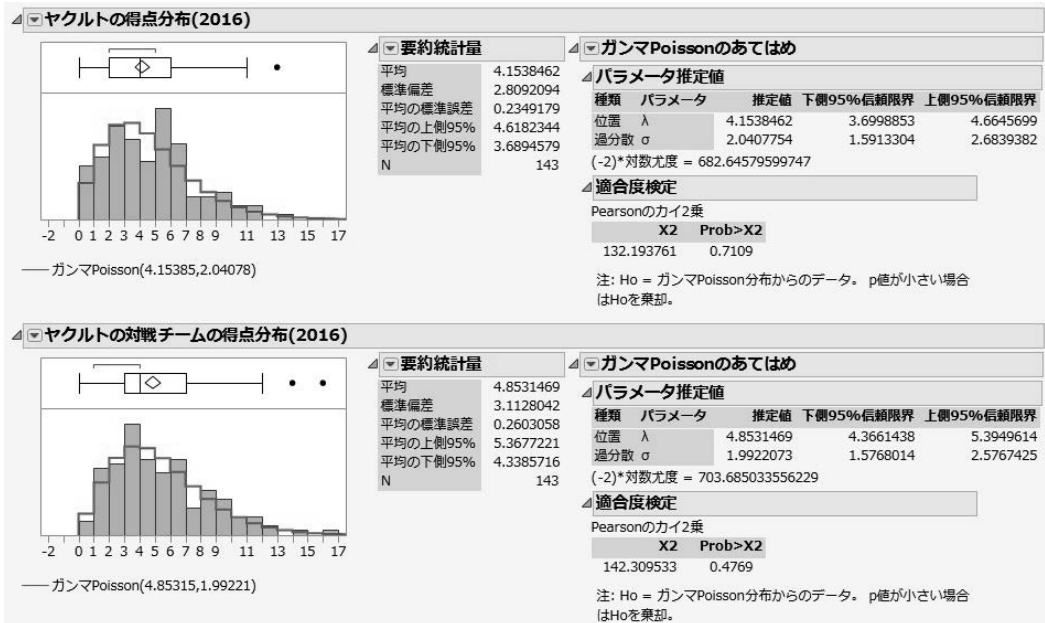
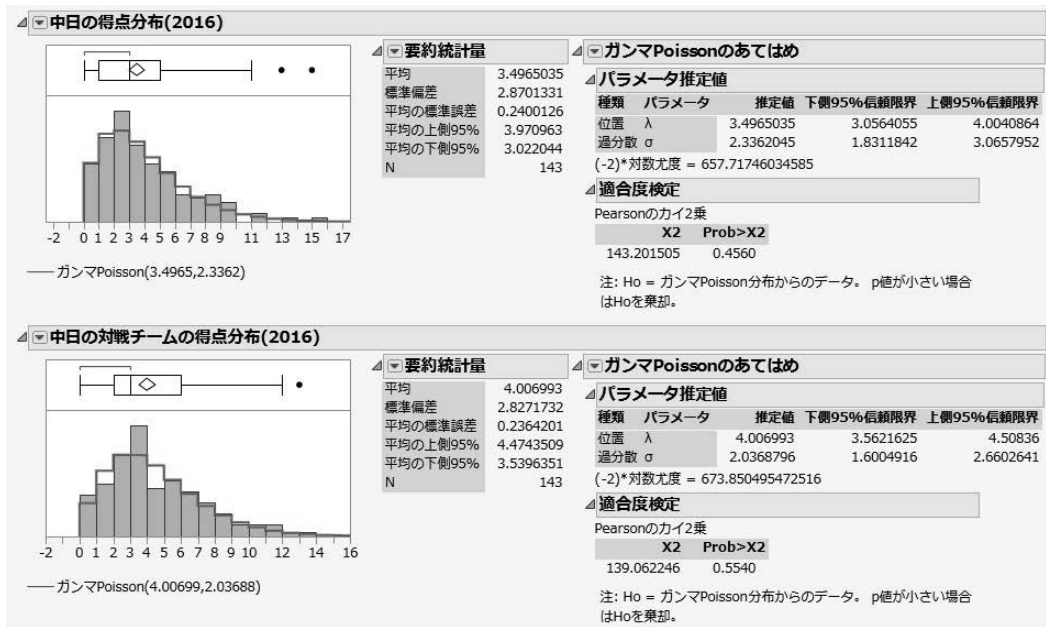


表9 セリーグ・ペナントレース結果(2016)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①広島	143	89	52	2	.631	-
②巨人	143	71	69	3	.507	17.5
③DeNA	143	69	71	3	.493	19.5
④阪神	143	64	76	3	.457	24.5
⑤ヤクルト	143	64	78	1	.451	25.5
⑥中日	143	58	82	3	.414	30.0

図63 6位:中日2016年シーズン



7.2 パリーグ 2016年シーズン

図65から図70を見れば分かるように、6球団の1試合あたりの得失点分布は、ポアソン・ガンマ分布による近似が良好であることを示している。

図64を見れば分かるように、このシーズンはソフトバンクと日本ハムの首位争いと、ロッテと西武のクライマックスシリーズのための3位争い、そしてオリックスと楽天の下位争いであった。結果は、優勝が日本ハムでソフトバンクは2位、3位は西武を大きく引き離れたロッテ、5位は楽天で最下位はオリックスだった。

図 64 パリーグ 2016 年シーズン平均得失点散布図

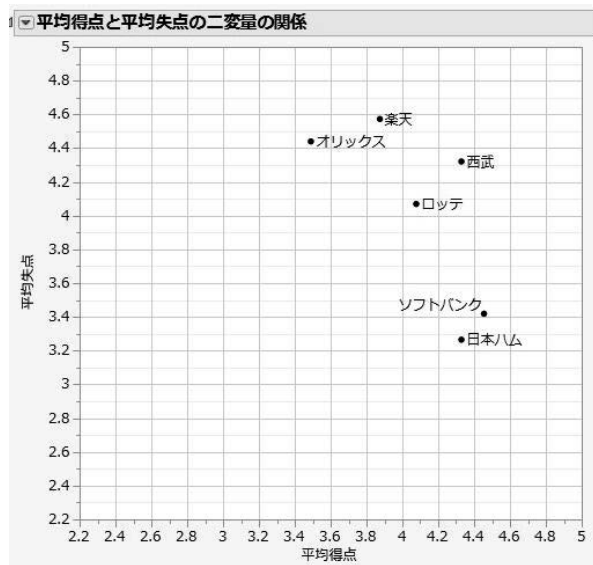


図 65 1位:日本ハム 2016 シーズン

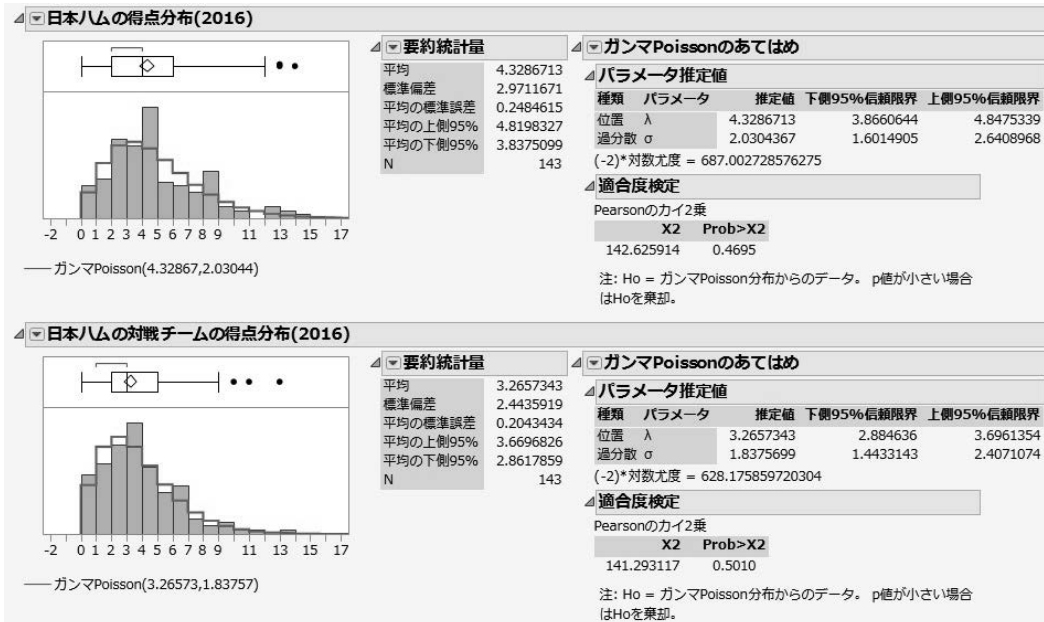


図 66 2位:ソフトバンク 2016年シーズン

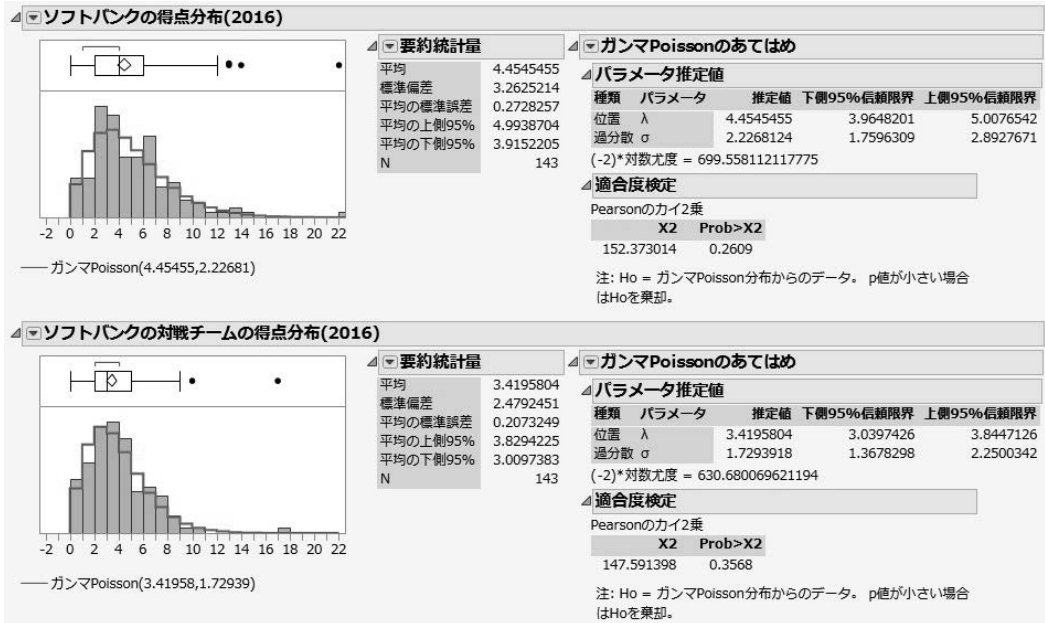


図 67 3位:ロッテ 2016年シーズン

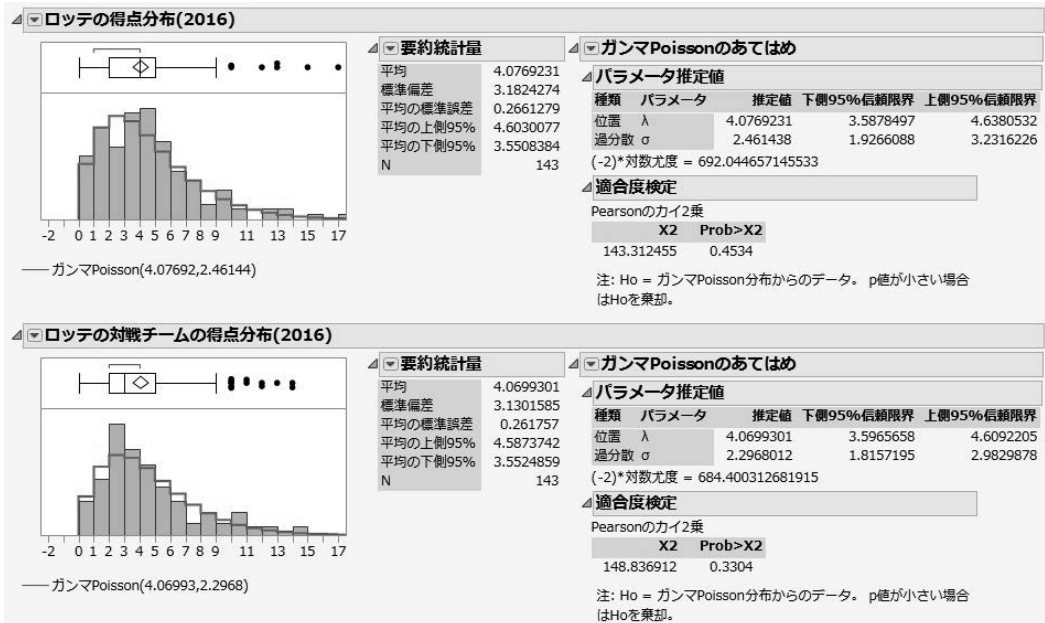


図 68 4位:西武 2016 年シーズン

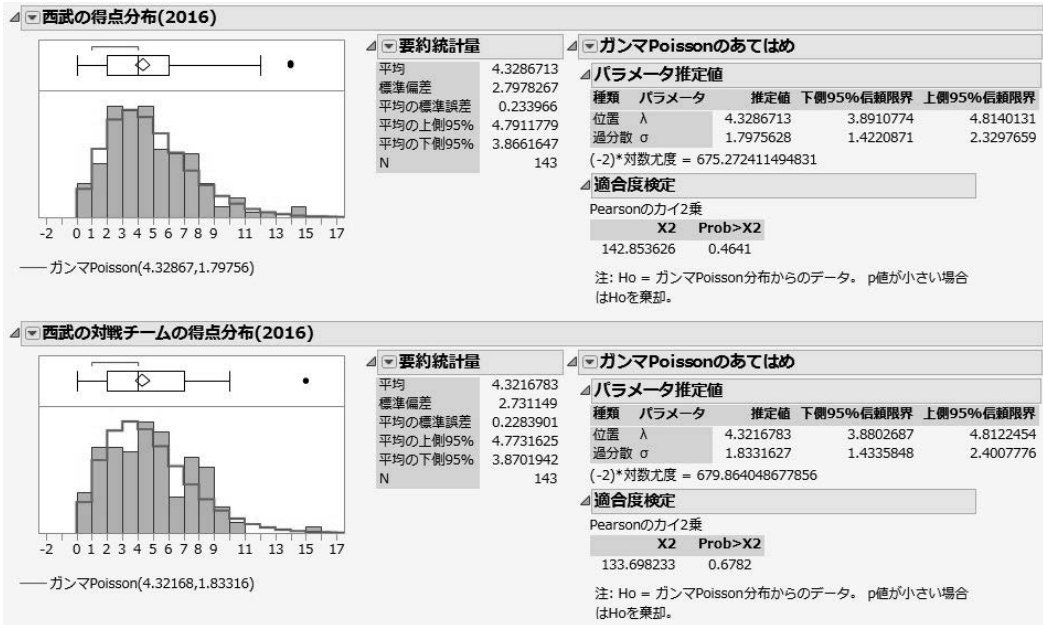


図 69 5位:楽天 2016 年シーズン

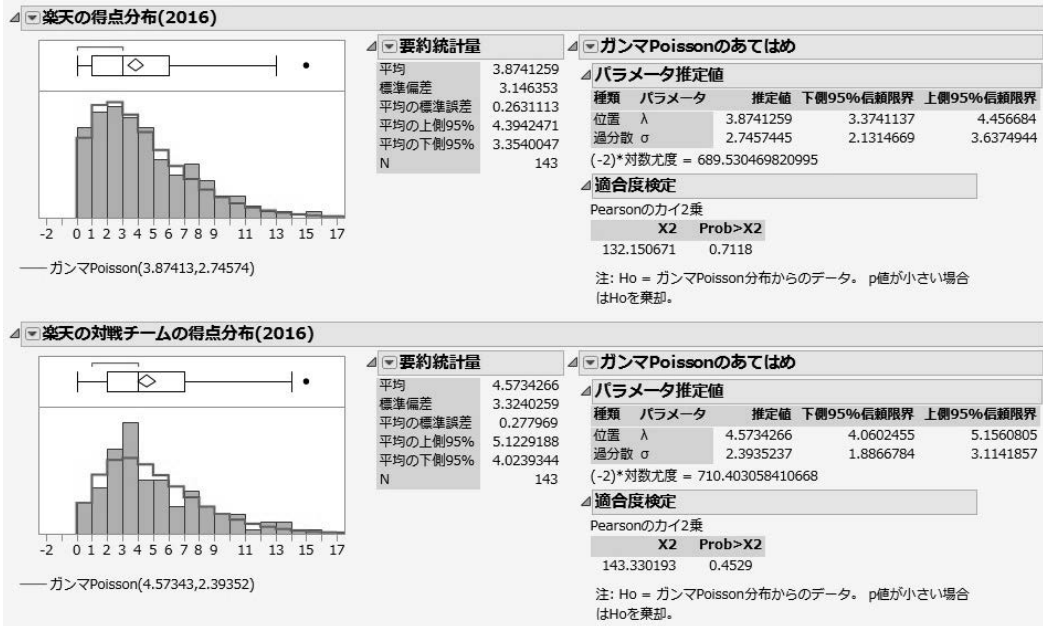
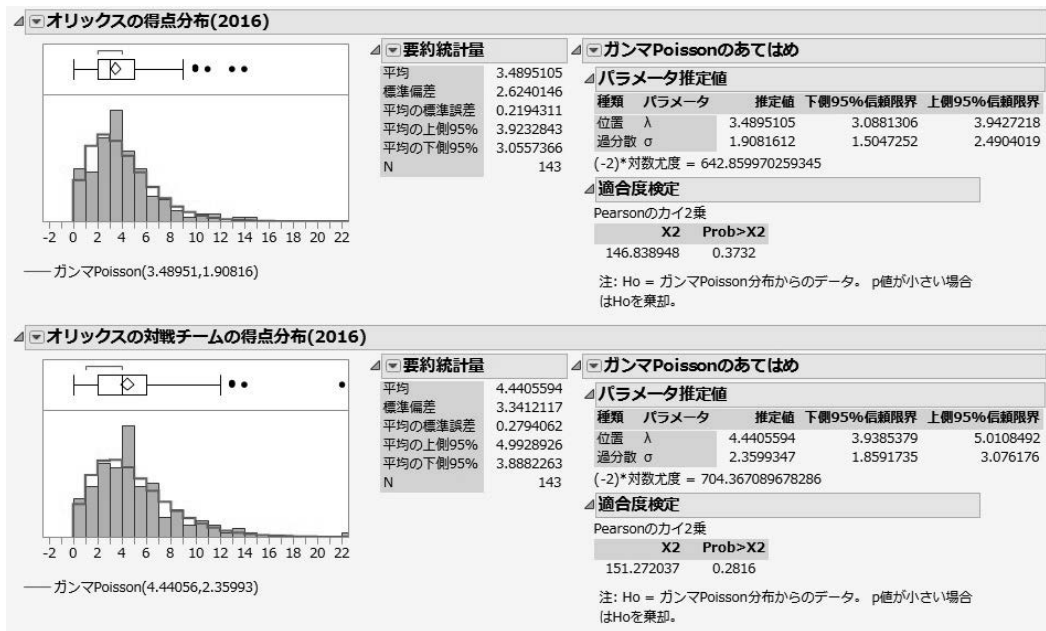


表 10 パリーグ・ペナントレース結果 (2016)

チーム名	試合数	勝利	敗戦	引分	勝率	ゲーム差
①日本ハム	143	87	53	3	.621	-
②ソフトバンク	143	83	54	6	.606	2.5
③ロッテ	143	72	68	3	.514	15.0
④西武	143	64	76	3	.457	23.0
⑤楽天	143	62	78	3	.443	25.0
⑥オリックス	143	57	83	3	.407	30.0

図 70 6位:オリックス 2016年シーズン



8 終りに

プロ野球の一試合あたりの得失点数の分布を調べるために、5年間のデータを入力し、各球団毎の得失点分布を求めた。その結果、一試合の得点分布も失点分布もガンマ・ポアソン分布で近似できることが分かった。このことは、ペナントレースのシミュレーションができることを意味する。各チームの母数を決定し、ガンマ・ポアソン分布に従うランダムな数列を発生させればよい。

あと一つ気づいたのが、得点効率の問題である。例えば、ロッテは散布図の位置のわりに上位を占めることが多いこと、ソフトバンクは反対に下位を占めることが多いように感じる。この問題に

については次の論文で明らかにしたいと思う。

参考文献

- [1] Dobson, A. J.(2002) *An introduction to Generalized Linear Models, 2nd edition*, Chapman and Hall (田中豊ら訳 (2008) 『一般化線形モデル入門 原著 第2版』)
- [2] ベースボール・レコード・ブック 2013 ベースボール・マガジン社
- [3] ベースボール・レコード・ブック 2014 ベースボール・マガジン社
- [4] ベースボール・レコード・ブック 2015 ベースボール・マガジン社
- [5] ベースボール・レコード・ブック 2016 ベースボール・マガジン社
- [6] ベースボール・レコード・ブック 2017 ベースボール・マガジン社