STAT 476 Project 2: Radiohead Sales Analysis

1. Executive Summary

In this project, 1995-1996 sales and airplay data for Radiohead's *The Bends* was provided for analysis. The aim was to build a strong model for fitting sales. Further covariates were obtained and constructed, including Radiohead's tours with R.E.M., releases of the singles "Fake Plastic Trees," "Just," and "Street Spirit," and the effects of the holiday season/Christmas, in order to improve the analysis.

Nine models were built and evaluated, including exponential-gamma (or Pareto Type II), Weibull, and Weibull-gamma regressions as well as latent class models. Each model was evaluated on BIC and MAPE, as well as on the interpretability of the model parameters. The final model selected was the exponential-gamma, which featured the lowest BIC and 3rd lowest MAPE of the group, while being the most parsimonious. However, a close runner-up was the 3-class Weibull-gamma model, which featured a sizable improvement in MAPE.

Model	LL	BIC	R^2	MAPE
Exponential-gamma	-1845.28	3741.93	0.86	17.03%
3-class Weibull-gamma	-1844.06	3775.05	0.88	15.58%

Finally, two applications of the models were demonstrated. First, by examining model parameters, a strong argument could be made that the peculiar sales pattern was in part due to lack of product awareness. Secondly, the model was able to estimate the sales impact of covariates such as particular singles releases. Both highlight the immense practicality and effectiveness of this approach to modeling timing data.

2. Initial Observations and Covariate Selection

First, we observe the sales and airplay data for the album (Figure 2.1, 2.2).

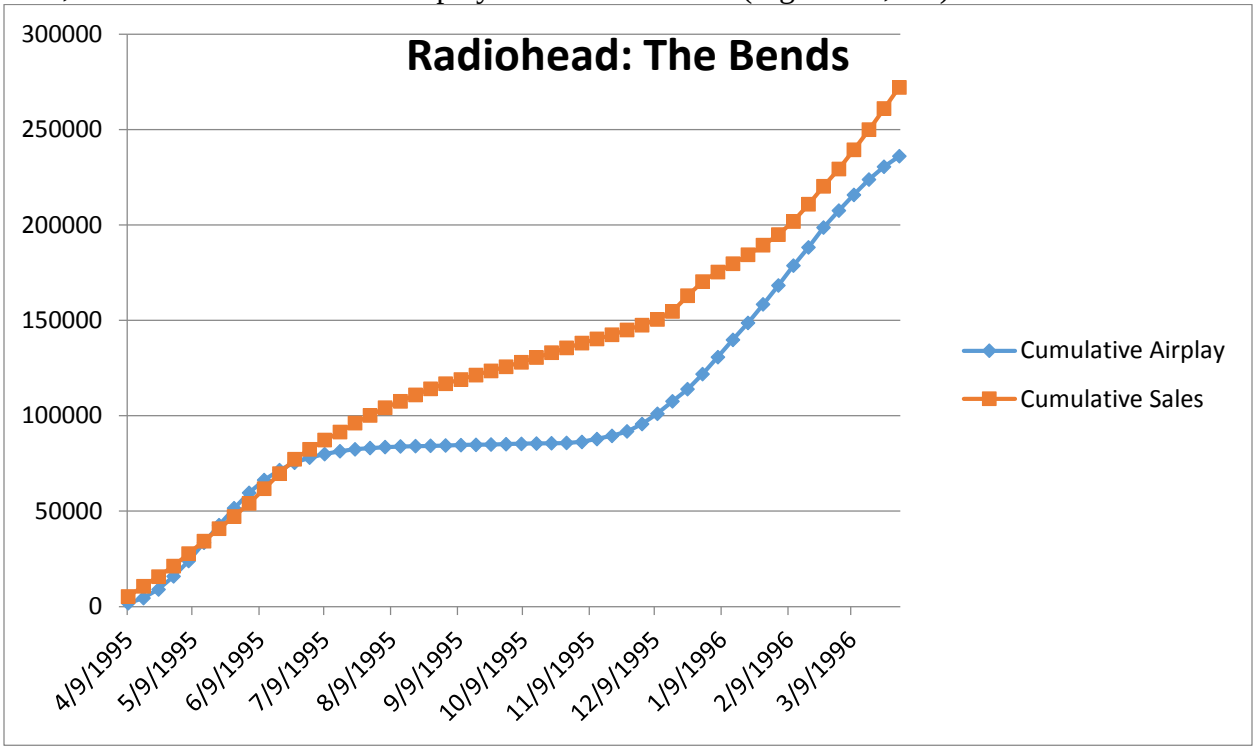

Figure 2.1 Cumulative sales numbers. Note the non-logarithmic curve, which is atypical of sales data.

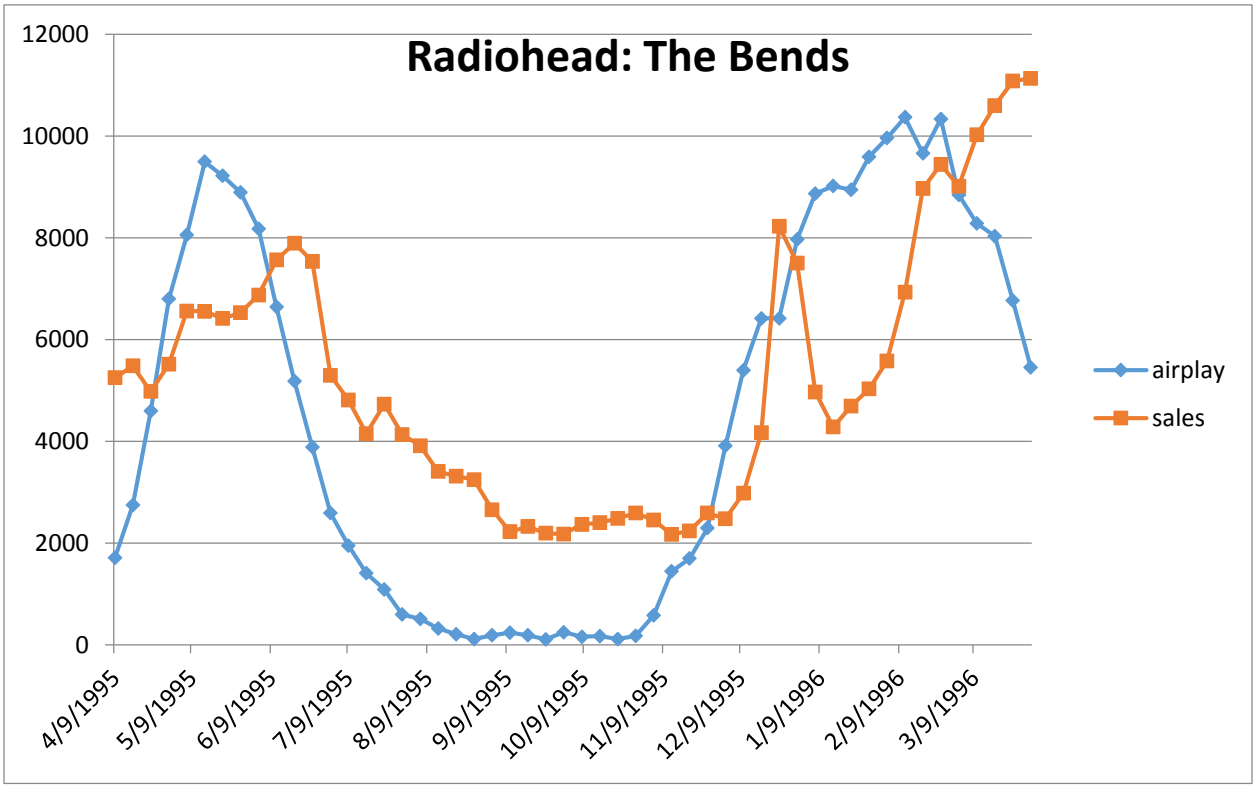

Figure 2.2 Incremental sales numbers. The irregularity is more pronounced, with the expected dip but an unexpected resurgence.

The cumulative sales initially follow a logarithmic curve, like most other sales data. However, surprisingly, there is a second inflection point around December 1995, where sales suddenly begin to pick up again. This increase in sales continues through March 1996, with no signs of slowing down. This phenomenon becomes very clear when looking at the incremental sales data, with abrupt spikes in sales in December and February 1996.

While it may be tempting to point to increased airplay as the driver of sales, logically this does not make sense. It is hard to explain why airplay would essentially cease from July 1995 to December 1995 before suddenly picking up. That many months after the album's release, why would an executive be confident that renewed airplay would revive sales? The most probable explanation is that some other event triggered both the increase in sales and in airplay.

From this reasoning, after some research, several predictors *other than airplay* were incorporated into the model building process:

1. Seasonality and the Christmas effect. This phenomenon should be taken into consideration for all sales data. For most products, sales typically increase dramatically during the Christmas season, and decrease during the week of Christmas itself. To incorporate these effects, two indicator variables were included, one for the month of December and one for the week of Christmas.
2. Singles releases. Three singles from the album were released in the timeframe of the provided data (five singles overall). These were "Fake Plastic Trees" on May 15th, 1995, "Just" on August 21st, 1995, and "Street Spirit" on January 22nd, 1995.[1] The release of a single would be expected to affect sales. However, their effect should also decay with time. To incorporate singles releases, two variables were constructed for each single release according to the specification $\beta d^{t-t_{release}}$, where β was the magnitude of the effect on sales and d the rate of decay.
3. Touring. In September, Radiohead played the opening set for 19 R.E.M. concerts in the U.S.[2] They had been supporting R.E.M. in other concerts overseas earlier that year; since this is U.S. sales data, those were disregarded. Touring would be expected to boost sales as well. The same specification as a single release was used to incorporate this effect.

A final important note is that in the models, all the sales are considered as first-time sales. Ideally the sales data would be separated into first-time and repeat sales, as the processes of trial and of repeat-purchase are very different, but that data is not provided. Still, since the album was just released at the start of the dataset timeframe, this should be a relatively safe assumption. Anyway, most people tend not to buy multiple copies of an album.

3. Model Building and Selection

Nine models were built for evaluation that incorporated covariate effects. They are:

1. Exponential-gamma. This model assumes the individual time to first purchase, denoted *T*, is exponentially-distributed with parameter λ, which characterizes an individuals' rate of trial (think of it as likelihood of making the purchase). λ is assumed to be gamma-distributed with parameters r, α, which allows for individuals to differ in λ.
2. Exponential-gamma with never-triers. This is the previous model but with the secondary assumption that out of the total relevant population of 2,000,000, only some fraction will ever even consider purchasing.

[1] https://en.wikipedia.org/wiki/The_Bends
[2] http://members.iinet.net.au/~darryl74/1995.html

3. Exponential-gamma with 2 segments. This assumes there are two subpopulations whose λ's are distributed with different parameters. It generalizes the above model.
4. Exponential-gamma with 3 segments.
5. Weibull-gamma. A major problem with the EG model is its memoryless property. That is, given that purchase hasn't occurred by time *t*, intuitively the probability of making a purchase between *t* and *t'* changes. However, that probability is unaffected by the given information in the EG model. We avoid this in the Weibull-gamma.
   Here, *T* is Weibull-distributed with parameter λ, and λ is gamma-distributed.
6. Weibull-gamma with 2 segments.
7. Weibull-gamma with 3 segments.
8. Weibull with 2 segments. In this model, there are two subpopulations which have Weibull-distributed *T*'s, but in each subpopulation the λ is static.
9. Weibull with 3 segments.

The model parameters are included in Table 3.1, as are several evaluation criteria. Log likelihood, BIC, R^2, and MAPE are displayed.

	1	2	3	4	5	6	7	8	9
	EG_cov	EG_NT_cov	EG_cov_2	EG_cov_3	WG_cov	WG_cov_2	WG_cov_3	W_2_cov	W_3_cov
*r*_1	1322.33	27.75	89.36	50.68	150.95	96.95	0.0003146		
*r*_2			0.61	1E-06		0.0001	4.856E-05		
*r*_3				1E-06			10.80		
α_1	517681.77	10875.13	36233.23	20538.3	66544.22	42579.65	1.09E-06		
α_2			1.11	1.77		0.12	0.12		
α_3				0.28			6513.99		
λ_1								0.0004205	0.068
λ_2								0.0023	0.39
λ_3									0.000654
*c*_1					1.04	1.03	23.98	0.25	0.12
*c*_2						0.10	1.31	1.04	1.35
*c*_3							1.10		1.21
π_1	1	1.00	1.00	1.00	1	1.00	0.39	0.0001	5.47E-08
π_2		1E-05	1.26E-05	6.03E-22		1E-05	0.17	0.9999	0.00
π_3				8.07E-07			0.45		1.00
β_airplay	0.03	0.03	0.03	0.03	0.04	0.03	0.07	0.03	0.09
β_season	0.19	0.20	0.22	0.23	0.17	0.18	0.02	0.17	-0.46
β_christmas	0.61	0.61	0.61	0.61	0.61	0.60	1.03	0.62	1.13
β_fpt	-0.05	0.00	0.16	0.16	-0.05	-0.03	-0.34	0.00	-0.22
β_j	-0.37	-0.44	-0.42	-0.43	-0.46	-0.47	-0.25	-0.49	-0.82
β_ss	0.62	0.62	0.67	0.67	0.57	0.58	0.52	0.57	-1.01
β_REM	-0.57	-0.86	-0.82	-0.83	-0.89	-0.90	-0.78	-0.90	-1.01
*d*_fpt	0.19	0.000001	0.87	0.87	0.31	0.04	0.00002	0.44	0.54
*d*_j	0.95	0.37	0.40	0.38	0.40	0.39	0.97	0.40	0.19
*d*_ss	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.58
*d*_REM	0.95	0.95	0.94	0.94	0.95	0.95	0.97	0.95	0.95
LL	-1845.28	-1845.34	-1845.21	-1845.22	-1845.25	-1845.26	-1844.06	-1845.26	-1842.61
BIC	3741.93	3746.00	3753.64	3765.51	3745.82	3761.64	3775.05	3753.74	3760.29
R^2_Incremental	0.86	0.86	0.87	0.87	0.87	0.86	0.88	0.87	0.93
MAPE_Incremental	17.03	17.29	17.42	17.46	17.54	17.83	15.58	17.64	13.89

Table 3.1 Model parameters and criteria

To choose a final model, both BIC and MAPE are used, the latter as a measure of goodness of fit.

Note that the Solver forced π to be as close as possible to 1 in EG_NT. That is, the model is rejecting the assumption that there is a subpopulation of individuals who will never buy the

album. The same can be said for EG with 2 segments and EG with 3 segments. Looking at the likelihood ratio tests, as well as the difference in MAPE, there is no good argument for the inclusion of segments. In fact, adding segments actually increases MAPE.

	test stat	df	p_value
LRT_1_2	0.119	1	0.731
LRT_1_3	0.136	3	0.987
LRT_1_4	0.127	6	1.000

The models also reject the segmentation in the case of Weibull-gamma. Notably, the 3 segment model actually has improved MAPE by ~2%. Still, considering that it has the highest value of BIC, the likelihood ratio test value, and that this marginal improvement costs 8 degrees of freedom, this model does have its cons.

	test stat	df	p_value
LRT_5_6	0.01119452	4	1.000
LRT_5_7	2.3857011	8	0.967

Finally we evaluate the latent class Weibull models. The 3-segment model actually features the best fit. However, considering the models have again forced π to 0, this model should be discarded. Furthermore, looking at the regression coefficients, all the single releases supposedly had a negative effect on sales. This does not match with what is expected, providing another reason to reject it.

With all of these considerations in mind, the best choice appears to be the regular exponential-gamma. It features the 3rd lowest MAPE and the lowest BIC. It is also the most parsimonious model. However, the 3-segment Weibull-gamma would have to be the second choice. Despite having the highest BIC, its lower MAPE and interpretability of model parameters make it a strong candidate.

4. Model Results

Let's examine the two models in detail.

First, for the EG, note the extremely large value of the *r* parameter, indicating a very homogeneous distribution of λ parameters throughout the population. That is, essentially everyone in the population has an equal rate of purchase. In fact, it might be advisable to run an exponential model with covariates instead. This was not done because the interpretation of such a model – complete homogeneity across the population – does not match well with reality. At the very least, this model acknowledges that individuals are heterogeneous. The most likely explanation for the current result is that the Solver was trapped in some local minimum of the log-likelihood function.

For the WG_3, there are 3 reasonably large segments of the population – 39%, 17%, and 45%. The first two have low *r* values, indicating highly heterogeneous distributions of λ. The last and largest segment has a high *r*. This does not match up very well with the EG model, which has one segment with extremely high *r*; this may be evidence that the EG is stuck somewhere in the log-likelihood. We must also examine the *c* values. The first segment has a very high *c* of ~24, indicating that the probability of buying given it hasn't yet occurred increases dramatically with time. This suggests that this segment is a group of hardcore Radiohead fans. The other two have *c* close to 1, which would degenerate to the memoryless case. More on that later.

Secondly, the regression coefficients are examined. The discussion will hold for both the EG and WG_3 models:

1. Airplay has a positive relationship with sales. This is somewhat obvious; the more the album is being played on the radio, the more people will be aware of it and perhaps buy it. A big question here is, as alluded to previously, whether airplay drives sales or lags behind sales. The answer is likely a mix of both.
2. The holiday season has a positive relationship with sales. This makes sense; sales of most products should increase during the holidays due to discounts, promotions, gifting, etc.
3. Christmas has a positive relationship with sales. This is surprising, since during the actual week of Christmas, people tend not to leave their homes to buy things. Looking at the data, on the week of Christmas, as well as the following week, there was a pronounced spike in sales.

   There could be two explanations. One is that the album was promoted during some holiday programming on radio or television. Another explanation would be that the week of Christmas, as recorded in the dataset, started on 12/24. Perhaps all the sales were due to regular retail activity, just in the latter half of the week, when holiday celebrations were over. Unfortunately, neither provides a compelling explanation for why sales spiked only temporarily.

   Adding covariates can only do so much for these kinds of models, but in the case of this abnormal behavior, suitable covariates could be very valuable.
4. Surprisingly, the releases of "Fake Plastic Trees" and "Just," holding all other variables constant, had negative effects on sales. What probably occurred is that the releases did not have any significant positive impact, and the natural decline in sales was attributed to the singles. The same holds for the coefficient of the R.E.M. tours, which was also negative. While it is theoretically conceivable that some potential buyers decided not to buy after those events, that story does not seem believable. Looking at the decay factors, the impact of "Fake Plastic Trees" decayed quickly, while that of "Just" decayed slowly. This can be explained by the fact that "Fake Plastic Trees" was dropped soon after the album; at this point the initial excitement from the album release had not yet subsided, so the release did not have much of an effect.
5. The release of "Street Spirit" had a positive relationship, holding all else equal, with sales. This is what we would expect from the release of a single. Doing some research, "Street Spirit" is one of Radiohead's most popular songs, so the fact that this release had a positive impact vs. "Fake Plastic Trees" and "Just" makes sense. Furthermore, the decay factor is high, indicating that sales would be increased for some time due to the release.

Finally, let's evaluate model performance. The cumulative and incremental sales projections are shown (Figures 4.1-4).

The models do an incredible job of fitting the cumulative sales figures. The expected sales fits are less impressive, but still manages to capture most of the nuances, only being thrown off by unexpected peaks in sales. Note that the sales numbers are divided by 1000.

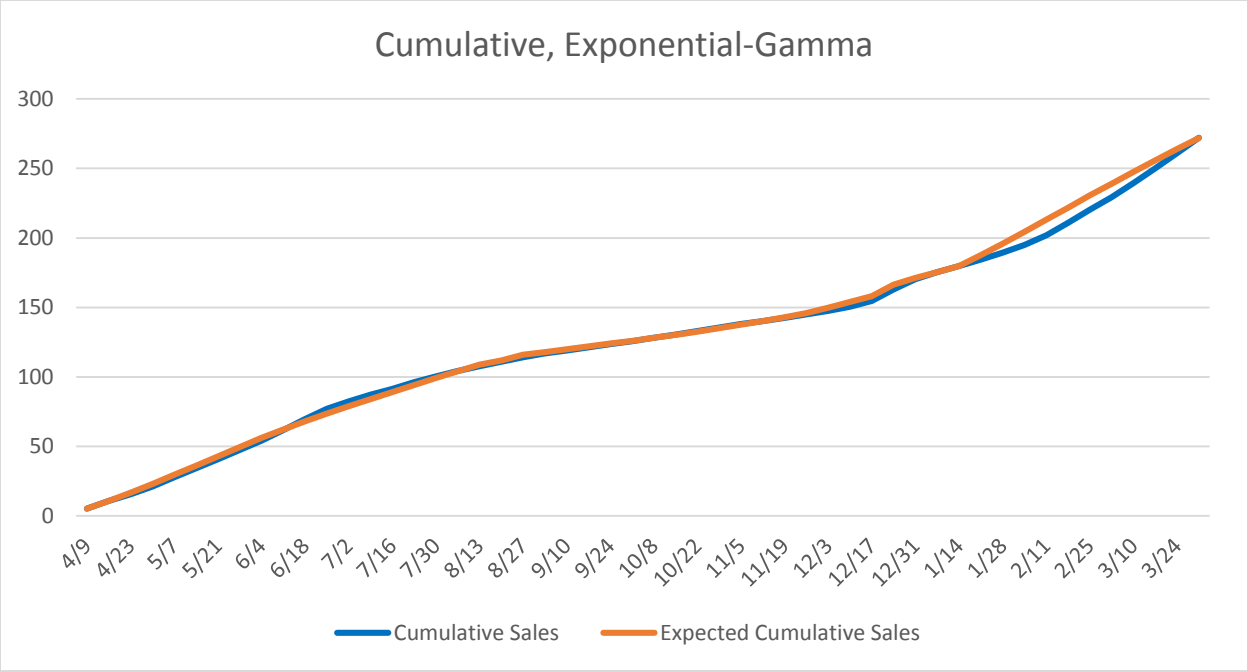

Figure 4.1 Cumulative sales fit for EG.

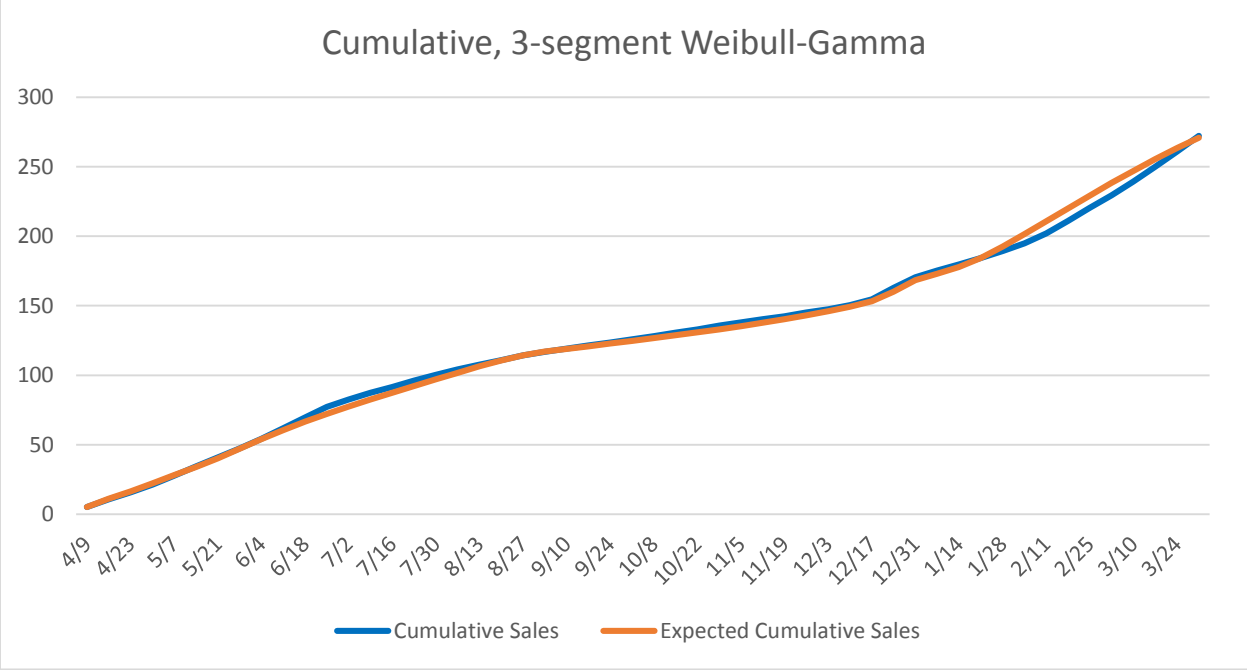

Figure 4.2 Cumulative sales fit for WG_3. They are hardly distinguishable.

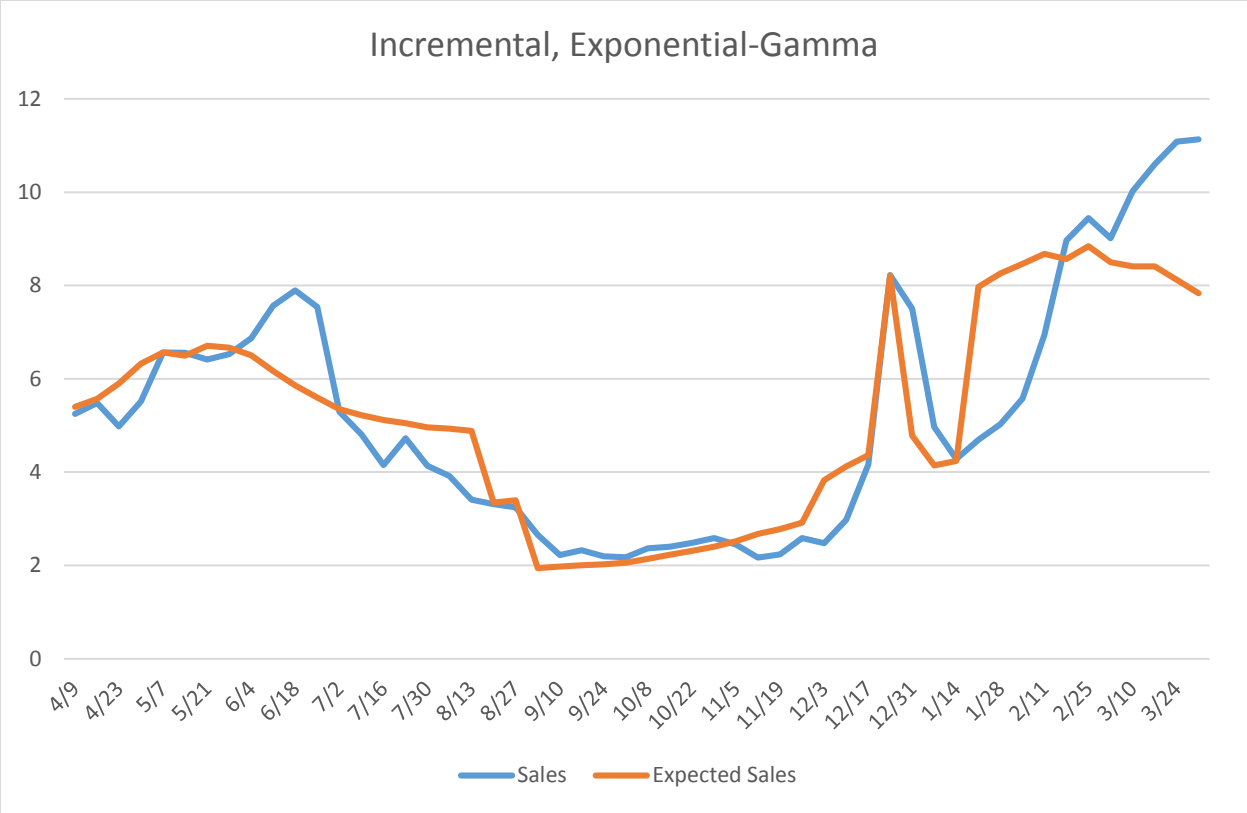

Figure 4.3 Incremental sales fit for EG. Less impressive than cumulative but still a good result.

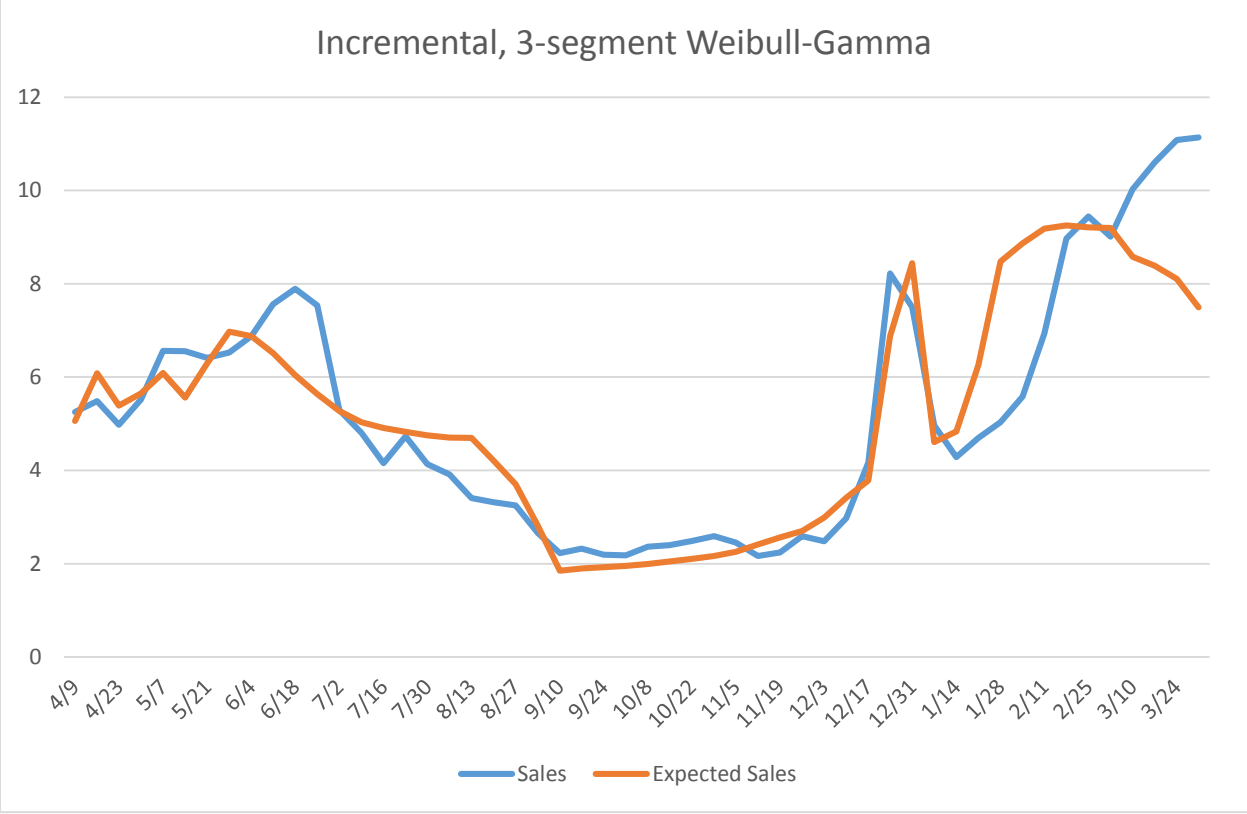

Figure 4.4 Incremental sales fit for WG_3. A slightly better fit, particularly in December.

5. Conclusions

What are some of the managerially relevant conclusions we can draw from the model? One major implication of the exponential-gamma model has to do with the aforementioned memoryless property of the exponential distribution. Things like probability of making a purchase given that it hasn't yet occurred should not be memoryless; in this case one would expect the probability to decline. The final model being exponential is very curious. This is further amplified by the fact that in all of the fitted Weibull models, including the 3-segment Weibull-gamma, most of the population have *c* of nearly 1.

The interpretation is that awareness of the album is low. This explains why for most individuals, the probability of trial given they haven't yet purchased the album is unaffected by time. If they haven't heard of the album, no matter how much time has passed, the probability won't change. This may be due to the band's popularity in the U.S. being low at the time (this was only their 2nd album), or the label failing to advertise properly, or even the way information was distributed during that era (it was basically the start of the Internet at the time). Nevertheless, this has direct business implications, especially since the renewal in sales 10 months after release shows that better initial promotion could have led to much better initial sales figures.

Other conclusions that can be drawn are the impacts of each of the predictor variables on sales. By turning off variables (setting β's to 0) and comparing the sales projections, the impact of the release of "Street Spirit," for instance, can be examined. A figure is displayed, using the final Exponential-Gamma model (Figure 5.1, 5.2). The caveat is that this methodology assumes that other variables do not come into play; in other words, that the increase in sales at the time of the release of "Street Spirit" can be entirely attributed to "Street Spirit."

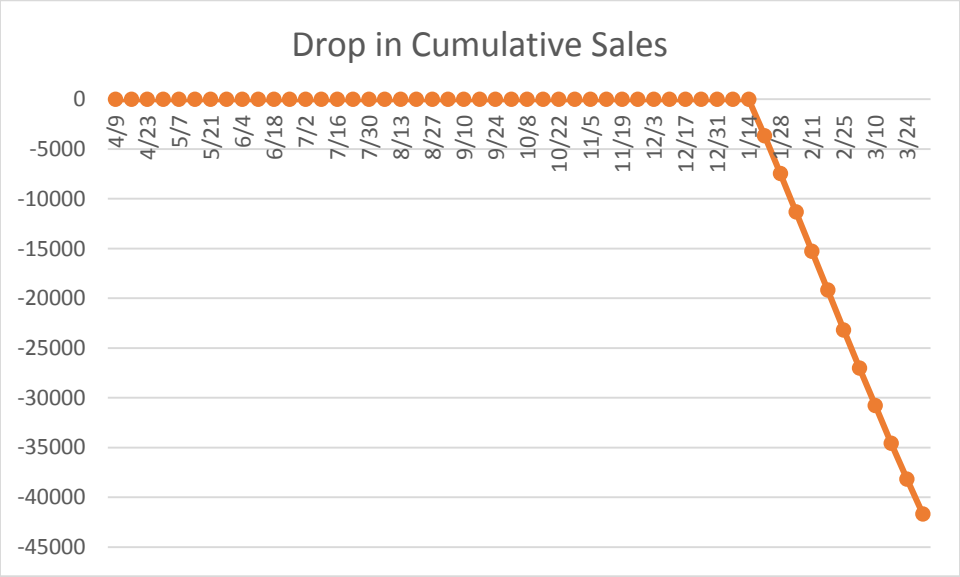

Figure 5.1

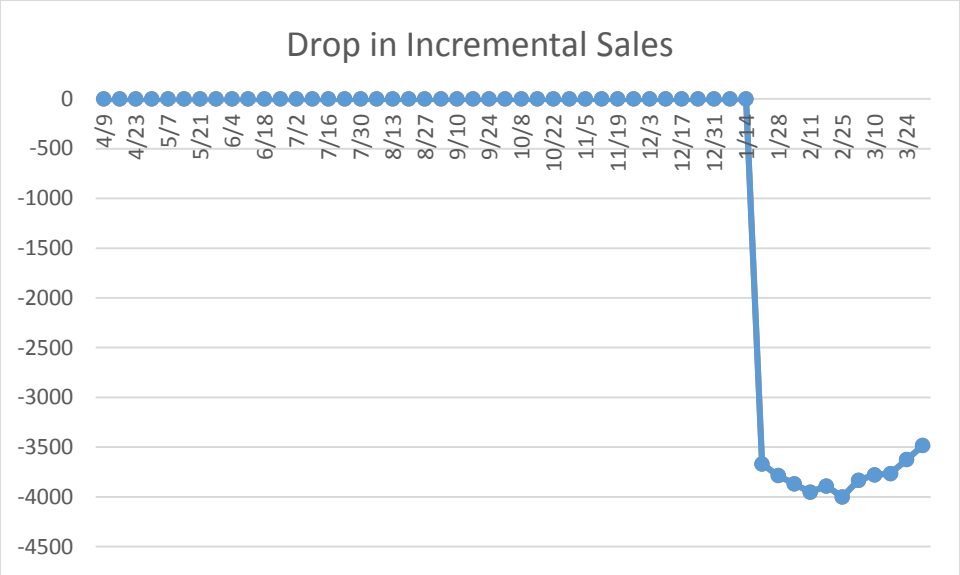
Drop in Incremental Sales
0
-500
-1000
-1500
-2000
-2500
-3000
-3500
-4000
-4500
4/9
4/23
5/7
5/21
6/4
6/18
7/2
7/16
7/30
8/13
8/27
9/10
9/24
10/8
10/22
11/5
11/19
12/3
12/17
12/31
1/14
1/28
2/11
2/25
3/10
3/24

Figure 5.2