Radiohead on the Radio?
Modeling the Sale of the Bends

April 6, 2016

Abstract

We construct different timing models for the sale of Radiohead's album, the Bends. We find that mixing distributions are not necessary. There is significant evidence for the existence of two classes of buyers, fans and the general public. We introduce holiday seasonality effects, tours and airplay as covariates acting on the hazard rate and find this can help explain purchasing behavior during particular periods. We derive the surprising conclusion from our final two segment Weibull with covariates model that only airplay compels the general public to purchase the Bends.

Introduction

We commemorate the anniversary of the Radiohead's the Bend by building probability models. Our goal is to construct a model that tells a compelling story while staying parsimonious (aka low model complexity). We will avoid truncated models as inferences on effective N are highly unstable. Our spreadsheet software is already extremely shaky when maximizing likelihood — we should be highly suspicious of any effective N inferred through models.

Data

Our data consists of the sales of The Bends over a 52 week period starting with the week ended April 9th, 1995. We also use airplay data (a weighted radio exposure metric), touring data, and holiday effects as covariates.

From examination of the sales data, it's clear that there are two different classes of customers. One are fans of Radiohead who purchased the album close to when it was released. The other are the public, whose purchase of the Bend likely depends on exposure to the album through the radio and tours.

Overview of Models

Table 1: Summary of Model Performance

Models	LLN	BIC	MAPE	# Parameters
Exp.	-1872641	3745298	11.81%	1
2 Seg. Exp.	-1872576	3745195	12.2%	3
EG	-1872642	3745312	11.81%	2
2 Seg. EG	-1872642	3745356	11.84%	5
W	-1872212	3744455	13.5%	2
2 Seg. W	-1843766	3687604	2.85%	5
2 Seg. W + A	-1843540	3687196	2.89%	7
2 Seg. W + A + H	-1842140	3684411	3.75%	9
2 Seg. W + A + H + T	-1841960	3684079	4.15%	11
WG	-1872213	3744469	13.5%	3
2 Seg. WG	-1844130	3688362	3.08%	7
W + A	-1855320	3710685	5.91%	3
W + A + H	-1855148	3710355	5.98%	4

Our main models are those in bold. "A" stands for airplay, "H" stands for holiday and "T" stands for touring. "H" and "T" are both indicator variables. "H" is 1 for 2 weeks beginning the middle of December, covering the weeks surrounding Christmas. The week of Christmas has an increased weight of 1.6. "T" is an indicator variable that is 1 if Radiohead was touring at the time.

We started with the most simple models, with parsimony in mind — we want to avoid complexity as much as possible. Therefore, building on the exponential and the Weibull, we introduced latent classes. Where those lacked in explanatory power, we introduced covariates to capture exogenous effects.

Our criteria for judging is an emphasis on interpretability of the model, then examination of the tracking plots, and finally BIC and MAPE. If we cannot interpret our model, we have not learnt anything about the purchase behavior. Then, if the tracking plots are off or exhibit aberrant behavior, we must come up with a reason for that deviation. Finally, BIC and MAPE provide quantitative measures that aid us in considering the best model that manages the conflict between fit and model complexity.

Our model provides substantial evidence that airplay, "A", is an extremely influentially covariate, which we will discuss in the later section on Weibull with airplay and holiday covariates. And the inclusion of airplay is easily justified: in the 90s, the publics' first exposure to a song is usually through the radio. If they liked it, presumably they will purchase the album.

Holiday effects, "H" are justified by figure 1. 21% of all sales of albums during the year of 2014 in the UK occurred in December; the rest are distributed close to uniformly across the other 11 months. That this is due to a seasonality effect is a reasonable assumption. We therefore allow for holiday covariates. As well, sales of albums are likely clustered close to Christmas. Therefore, during the week of Christmas, we allow a greater weight of 1.6 for the indicator variable.

Finally, consider touring data, "T". Concert goers are often exposed to new artists for the first time through opening acts for headliners. At major music festivals, Radiohead might have opened for more prominent artists, and thereby exposed many to their music for the first time. A brief search across their concerts in 1995 and 1996 shows that this is the case. [1] Notably, Radiohead took a break from touring after

[1] Radiohead Tours: http://www.greenplastic.com/gigography/index.php?year=1991-1992-1993-1994-1995-1996-1997-1998-1999-2000-2001-2002-2003-2004-2005-2006-2008-2009-2010-2011-2012

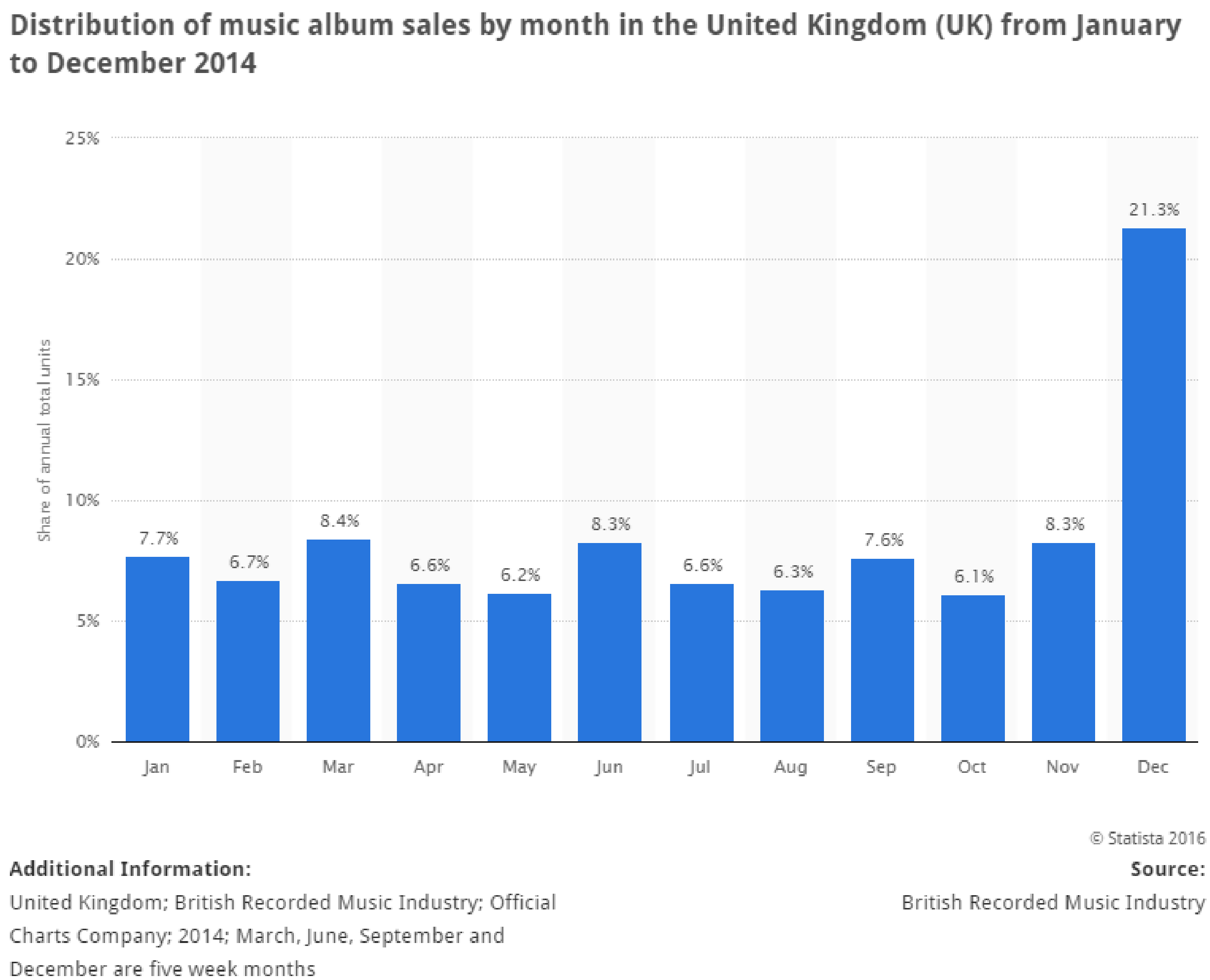

Figure 1: Abnormal Sales in December

December 18th, 1995, and resumed on March 14th, 1996.

In particular, not allowing for latent classes results in extremely poor fit. This is true even when we allow for heterogeneity through the gamma mixture. This is likely due to the extreme differences between the two classes of buyers, and that the gamma distribution only allows for one mode. As well, we find that latent class Weibull Gamma models are extremely difficult to optimize, and fail to provide greater explanatory power than the latent class Weibull, despite burning 2 more parameters.

Weibull with Airplay and Holiday Covariates

Table 2: Parameters for Weibull with Airplay and Holiday Covariates

	Specification
λ_0	0.001707
β_H	0.1128
β_A	96.32
c	0.9961

We start by considering the exception. Weibull with covariates does a surprisingly good job at capturing the cumulative sales of the Bend. This provides evidence against the existence of two separate classes of people. This model says that that there is only one class of consumers, who are heavily swayed by exposure through the radio and by Christmas effects. Indeed, our cumulative tracking plot shows that this model performs well up until the holiday season, at which point the two diverge. The end points for both are the same. This is in spite of the fact that we have included the Christmas season as a covariate.

Furthermore, that airplay exhibits a positive effect on the hazard rate is unsurprising. Greater radio exposure leads to more purchases in any one period. We note with interest that duration dependence is close to 1; this model may have failed to pick up on duration dependence due to the extreme differences between the two classes. Otherwise, it could be that this model heavily relies on airplay and holiday effects to

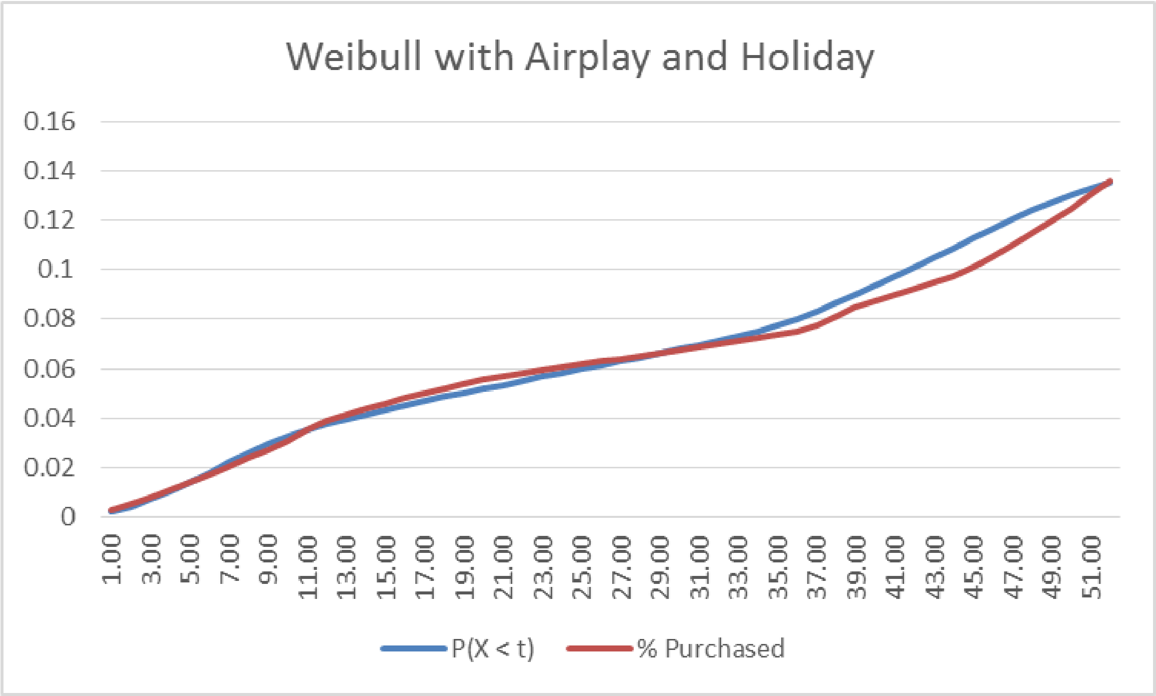

Figure 2: Cumulative Tracking Plot

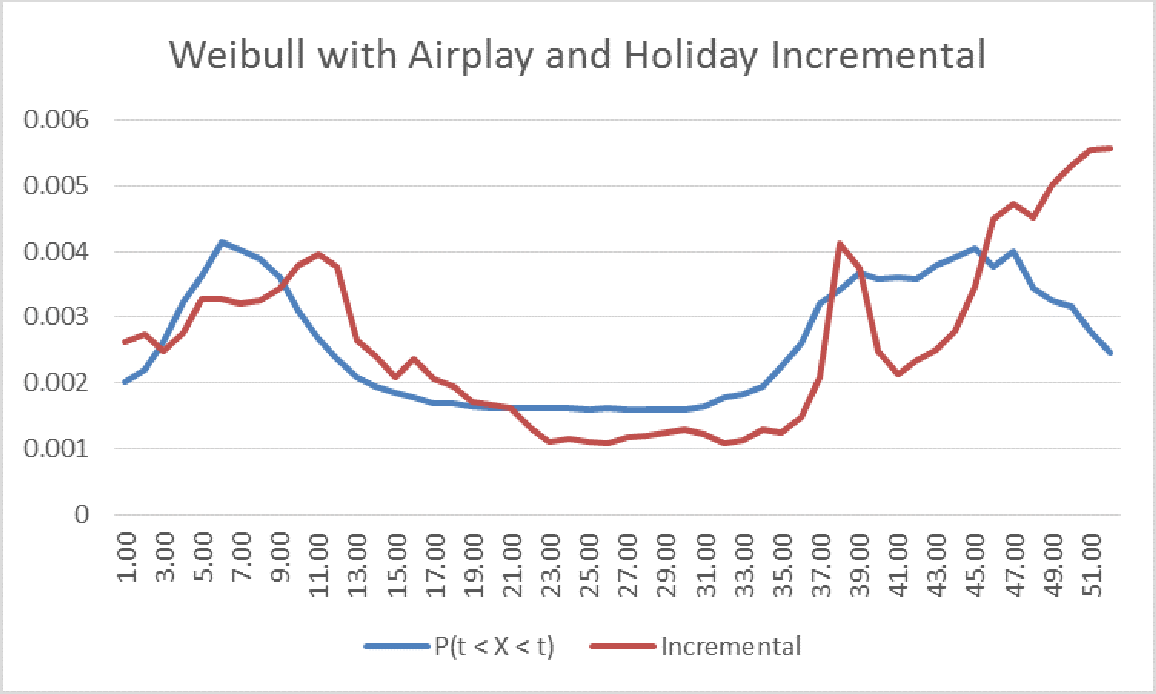

Figure 3: Incremental Tracking Plot

manipulate duration dependence.

For all the simplicity of this model, we are still puzzled by the divergence in the incremental tracking plot after the holiday season. Our model predicts continually declining sales, even as actual sales are climbing. This is likely due to the existence of a second class of buyers who are only now making purchases.

As well, there is a substantial difference in the LLN and BIC between this model and the two segment models we will explore next. We therefore abandon this approach in pursuit of greater explanatory power and a more compelling story.

Two Segment Weibull

Table 3: Parameters for Two Segment Weibull

	Segment 1	Segment 2
λ	0.02671	8.591E-10
c	1.446	4.067
Proportion	6.091%	93.91%

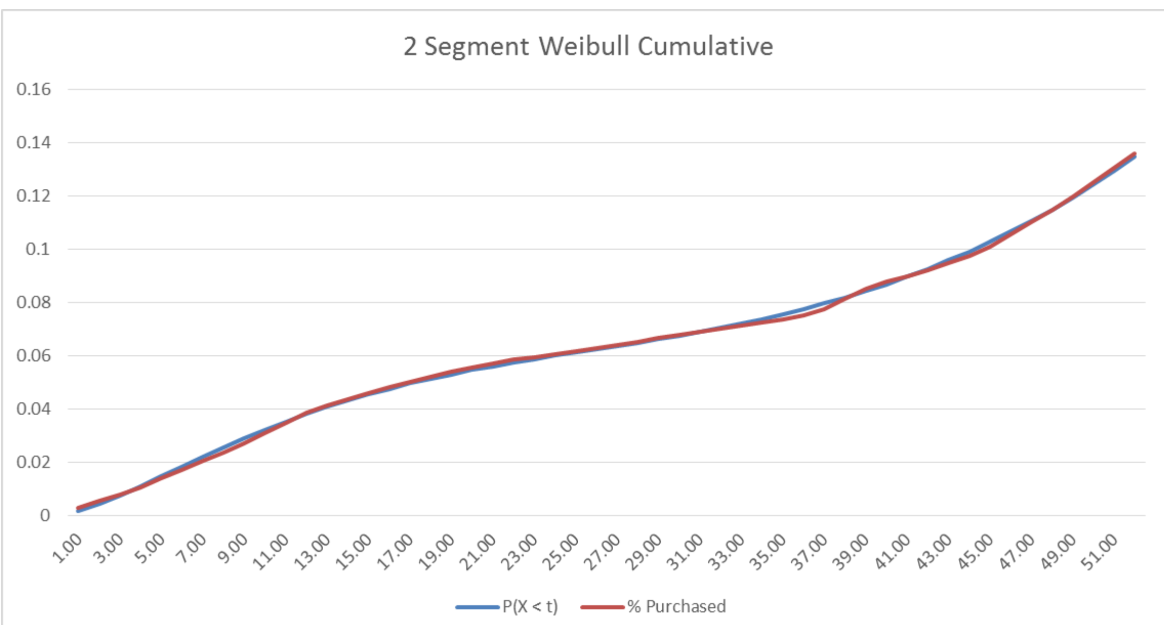

Figure 4: Cumulative Tracking Plot

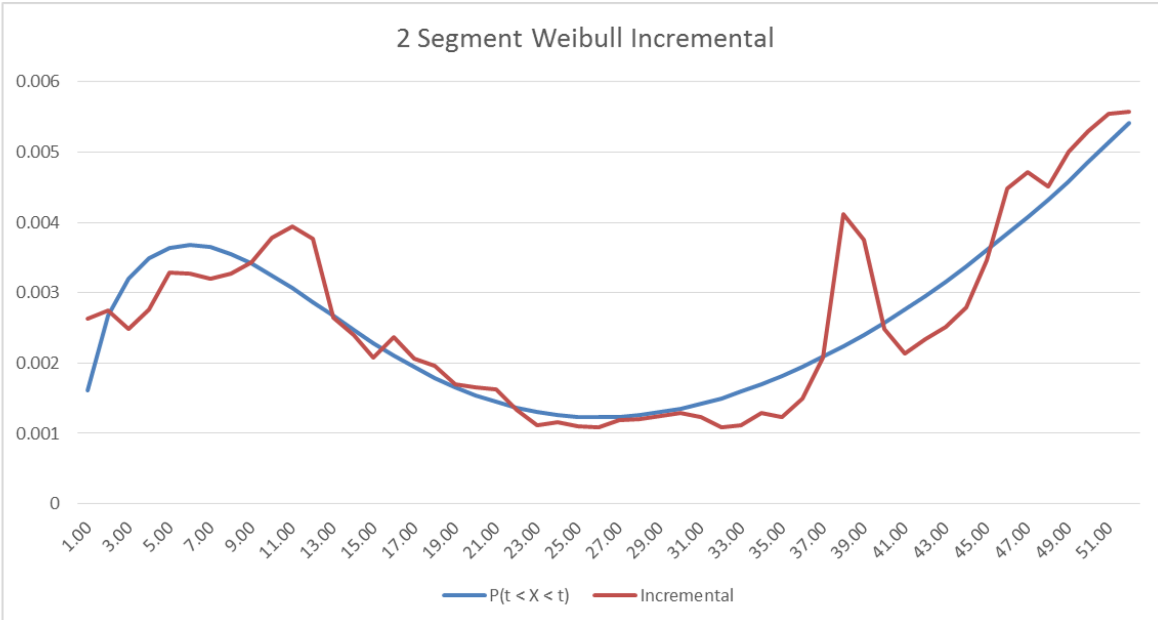

Figure 5: Incremental Tracking Plot

In this model, we allow for two different classes of people as discussed previously. Immediately, we see from the cumulative and incremental tracking plot that fit has improved drastically. The former follows the actual sales pattern extremely closely. That this has occurred without using gamma mixture is a plus — the more parsimonious the better. In the incremental plot, the divergence problem we found in the previous Weibull with covariates model has been addressed: our model rises along with actual sales.

This model has a compelling and intuitive story. Recall in the introduction that we suspected the existence of two distinct classes of people based on the sales data along. That suspicion is confirmed here. One segment has an extremely low λ along with a large $c_2 = 4.067$ duration dependence. This segment is likely the general public who are swayed by exposure to Radiohead through the radio or through concerts to purchase the Bend. This explains the positive duration dependence. The longer they wait, they more likely they are to have been exposed to Radiohead, and the more likely they are to make a purchase. The other segment has a λ that is six orders of magnitude larger; these are the fans of Radioead who are likely to make a purchase close to the release dates. Indeed, our model shows that 50% of fans will have made purchases by week 10, and 90% before week 22. They also exhibit positive duration dependence. Assuming fans of Radioheads know each other (likely due to the self-similarity effect in social networks), the longer they wait after the release of the Bends, the more they are convinced by word of mouth, concerts and radio exposure to purchase the bends. Furthermore, our model shows that 121,828 members of the population of 2,000,000 are fans. This makes sense as well — Radiohead was still relatively unknown at that time and only a small group will have been familiar with them. Interestingly, it appears that we do not need to explicitly allow airplay to affect duration dependence (hazard rate) to explain the sales pattern.

BIC and Log Likelihood also point to this being a better model. BIC indicates dramatic information gain over the previous Weibull with covariates model. This is even after penalizing for the additional parameter used to tease out the two segments. Furthermore, the increase in LLN shows a much improved fit. This is supported by the drop of the average absolute percentage error from 5.98% to 2.85%.

Without covariates however, the spike in actual sales around Christmas season is unexplained and this isn't visible on the cumulative tracking plot. It begs to be explained through covariates, and we oblige this implication to consider the two segment Weibull with covariates.

Final Model: Two Segment Weibull with Airplay, Holiday and Touring Covariates

All the advantages of interpretation of the two segment model (discussed previously) hold here. And this model, which is our chosen final model, resolves the divergence in incremental sales around Christmas. Allowing that people's purchase behavior change dramatically in December, we find that purchases for both segments (fans and general

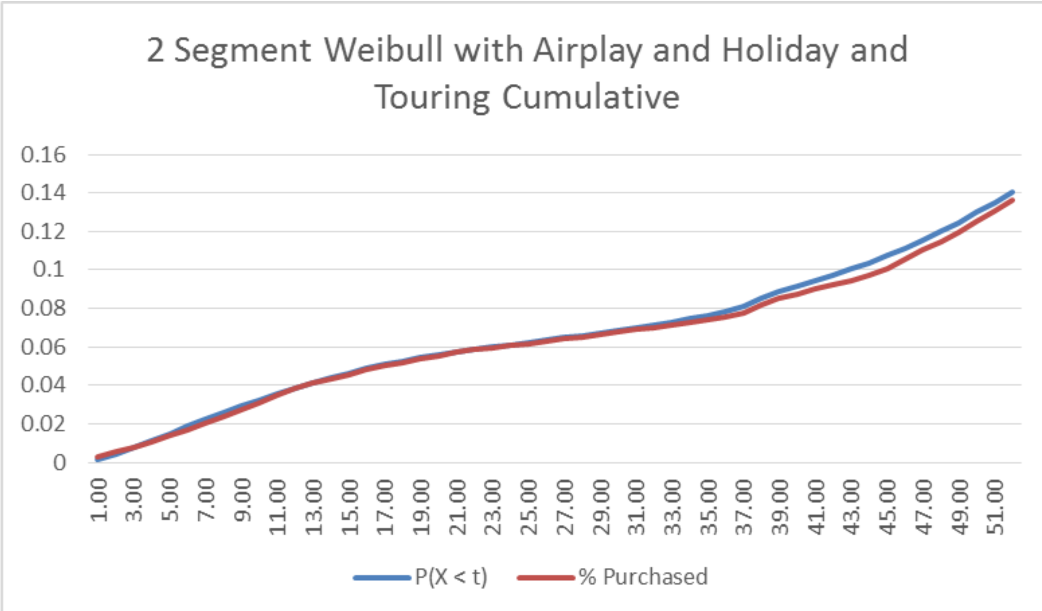

Figure 6: Cumulative Tracking Plot

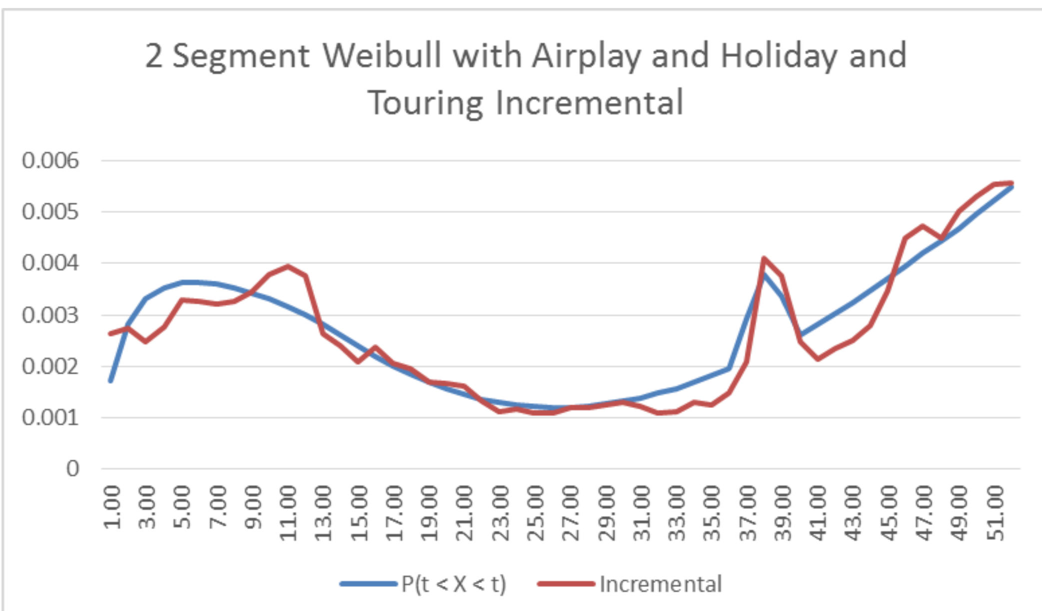

Figure 7: Incremental Tracking Plot

Table 4: Parameters for Two Segment Weibull with Covariates

	Segment 1	Segment 2
λ	0.02671	8.591E-10
c	1.446	4.067
β_A	-10.22	3.430
β_H	0.3270	0.3258
β_T	0.0740	0
Proportion	6.171%	93.83%

public) exhibit an uptick in line with the actual sales pattern. Interestingly, the holiday effect on both segments are similar (0.3270 vs. 0.3258), despite the fact almost all fans have purchased the Bend by that point.

For the fans, it appears that airplay has a negative multiplicative effect on their hazard rate. This is a surprising result. It's likely an artifact of the fans' propensity to purchase the Bend right away, and thus the model's finding that despite increased airplay that fans are not purchasing (again!). As well, touring has a positive multiplicative effect for fans; this may be explained as fans being more impatient to purchase the album after hearing Radiohead live.

For the general public, airplay has an unsurprising positive multiplicative effect on the hazard rate. Greater radio exposure leads to greater likelihood of purchase within any period. Interestingly, our model finds that tours have no effect on the hazard rate. This implies that the public do not attend rock concerts—and therefore could not have been exposed to Radiohead through this mechanism. This makes sense, as the type of person to attend rock concerts are most likely the ones who already know of Radiohead and likely fans as well. This attributes all conversions of the general public to radio exposure.

We pick this model as our final because of its compelling story and its ability to capture incremental sales behavior. BIC also shows significant information gain over the two segment Weibull alone; BIC improvements are in spite of the fact that this model uses 6 more parameters. LLN shows a better fit as well. However, MAPE has increased from 2.85% to 4.15%. We accept this as it isn't a dramatic increase, and all other signs point towards this being the better model.

The biggest problem with this model is that we are unable to solve for new estimates of λ and c for each segment. In 11 dimensions, Solver has an extremely difficult time navigating the likelihood space and finds that every solution close to the Two Segment Weibull model is a local optimum. We obtained our estimates by taking λ, c as given, and optimizing for βs and π alone.

Introducing the covariates may have aligned the incremental sales more, but cumulative sales has suffered, Beginning at the holiday season, our model begins to overestimate cumulative sales. This is likely due to the overestimation of incremental sales right before Christmas and right after. After that, our incremental sales are again in line with overall trends — hence our cumulative sales estimate runs parallel to the actual sales after Christmas.

Discussion & Applicability

We note that overfitting is a severe problem. Our models will likely generalize extremely poorly for any other albums by Radiohead, or for any other artists. This is because we only trained on the sales of one album for one artist — particularities of Radiohead purchasers (genre of music, income) and so on will not generalize to all music purchasers in general. Our models will make predictions that are likely only applicable for sales of other Radiohead albums.

Furthermore, given the extreme differences in sales data, we do not believe a holdout period for the sake of assessing out-of-sample performance makes sense here. It's clear that the pattern of behavior in the first months are due to fans, and the latter due to converted members of the general public. Therefore, we do not have an honest assessment of out-of-sample performance.

As well, there are patterns that we cannot explain. For example, there is no evidence of a depression of sales for albums in general in January. All incremental tracking plots show that we cannot explain the drop in sales after Christmas. As well, none of our models explain the sudden bump in sales around week 11 and week 46. These are failures of our models in spite of allowing for heterogeneity through c or λ and covariates.

A multicollinearity problem may affect the dependability of the Holiday indicator variable "H". We note that airplay for the Bends rises dramatically just as Christmas occurs. Thus, airplay only may already include the effects of Christmas through this high correlation.

Conclusion and Next Steps

The sale of the Bends by Radiohead is a timing process that can be modeled reasonably well without the use of mixture models. Our two segment Weibull model with covariates captures well the cumulative and incremental sales of the album, with a mean absolute percentage error of only 4.15%. We have seen evidence for two classes of buyers, fans and the general public, and that covariates (airplay, holiday seasonality effects and tours) have significant but unequal effects on the hazard rate. We also find the surprising result that airplay alone drives sales for the general public, not tours, which only positively affect the hazard rates of fans.

For next steps, we should analyses the sales of Radiohead's other albums, as well as the sales of other rock band's albums. This allows us to build a more reliable model for the purpose of forecasting the sales of rock bands in general. To obtain a characterization of music album sales in general, we would need to conduct a similar analysis across all different genres, a painstaking endeavor.

Otherwise, if we were only concerned with forecasting the behavior of sales of Radiohead's albums, we should forecast over the next year and test this against actual sales. This gives us an honest assessment of the performance of our forecasting model.

Appendix

Below are tracking plots for all other models we've examined.

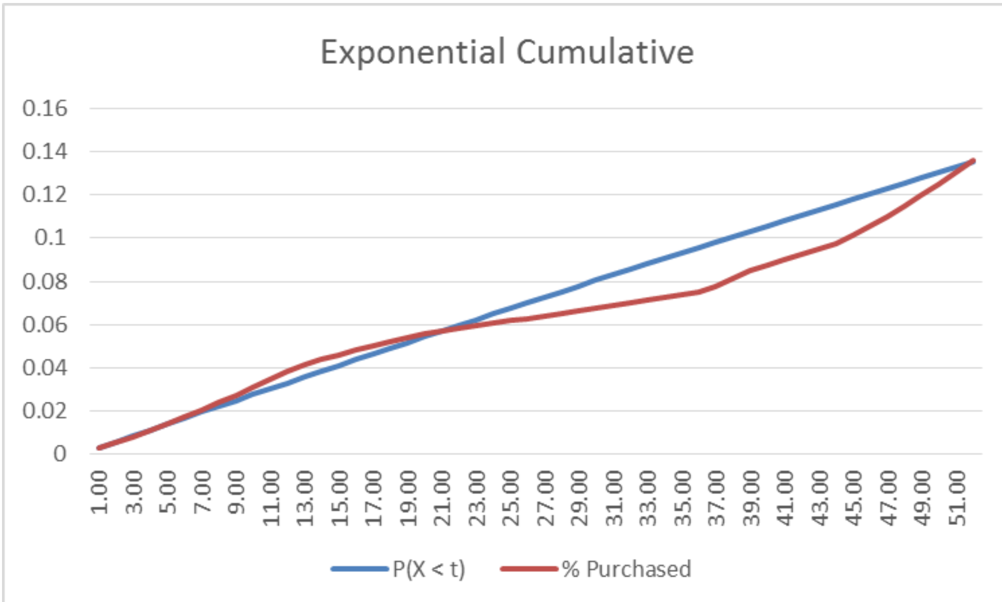

Figure 8: Cumulative Tracking Plot

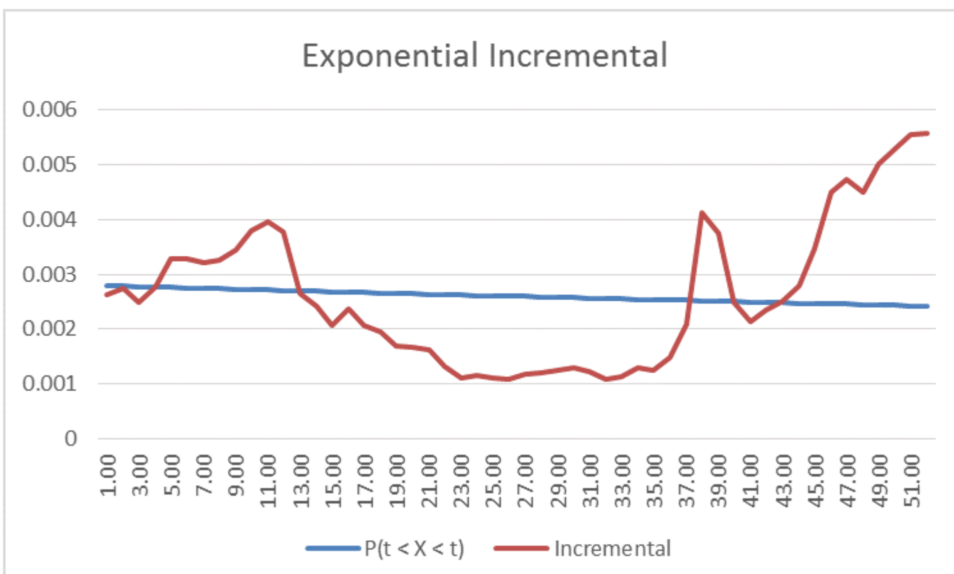

Figure 9: Incremental Tracking Plot

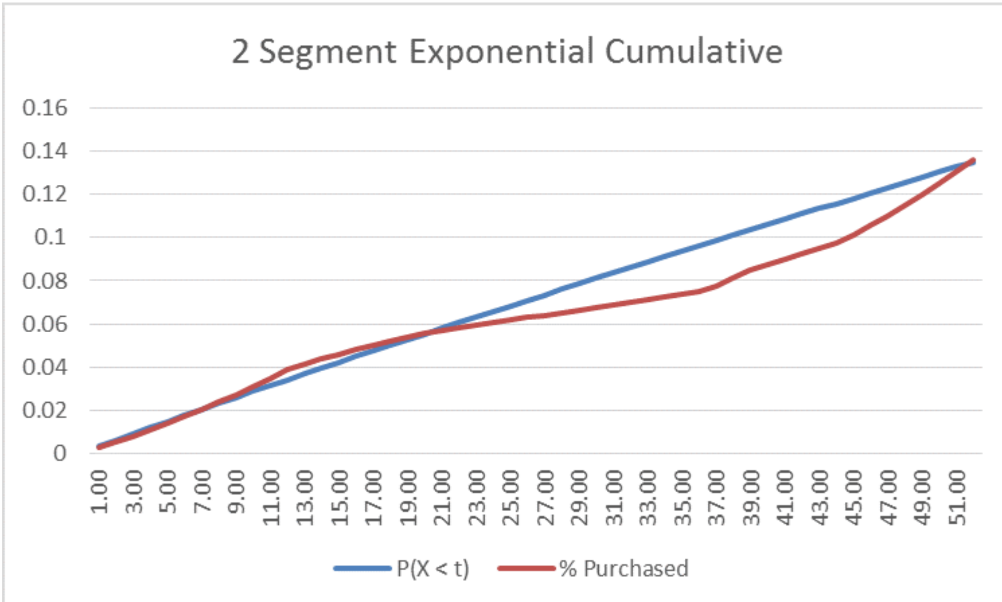

Figure 10: Cumulative Tracking Plot

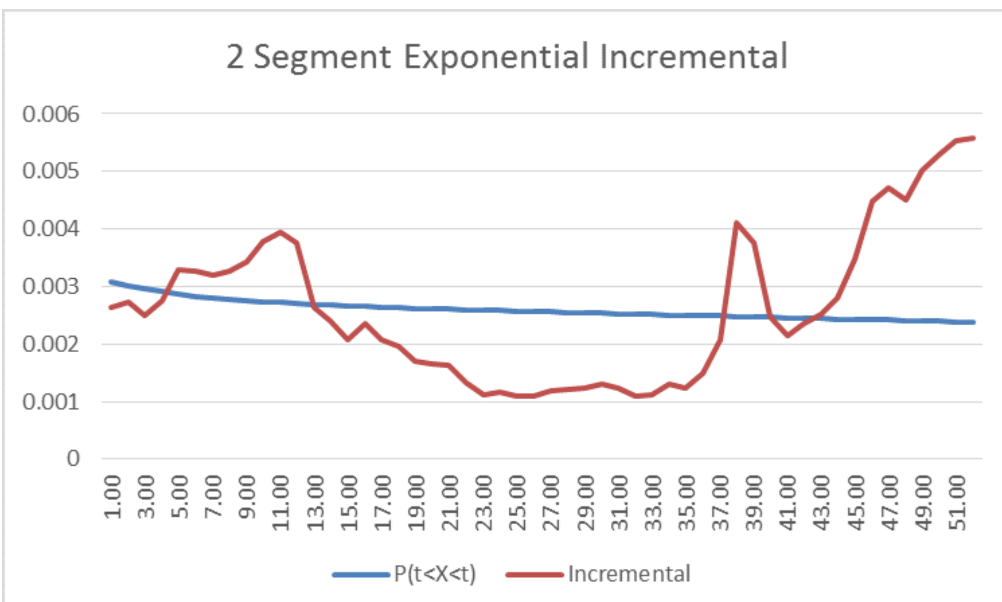

Figure 11: Incremental Tracking Plot

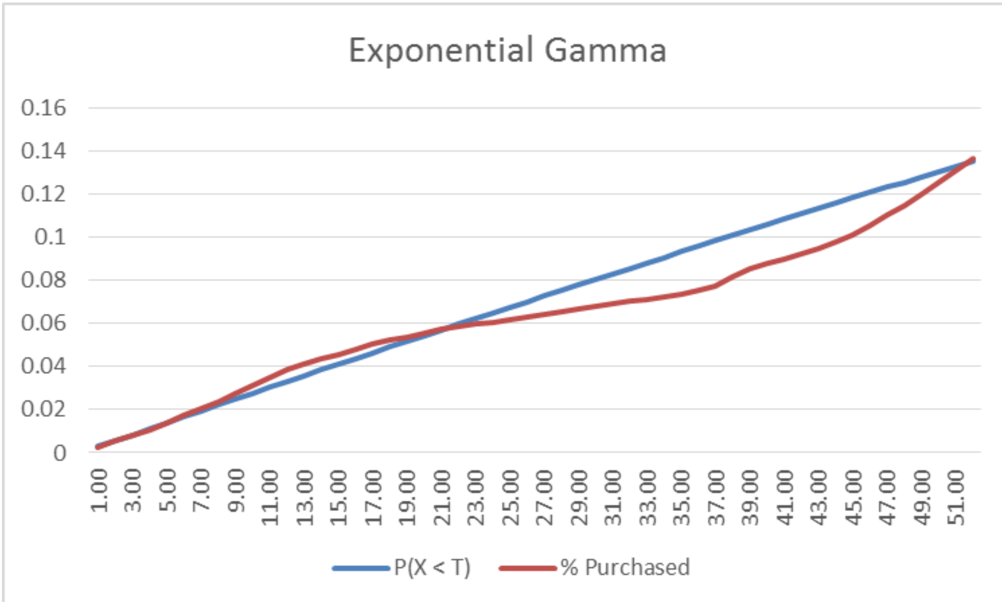

Figure 12: Cumulative Tracking Plot

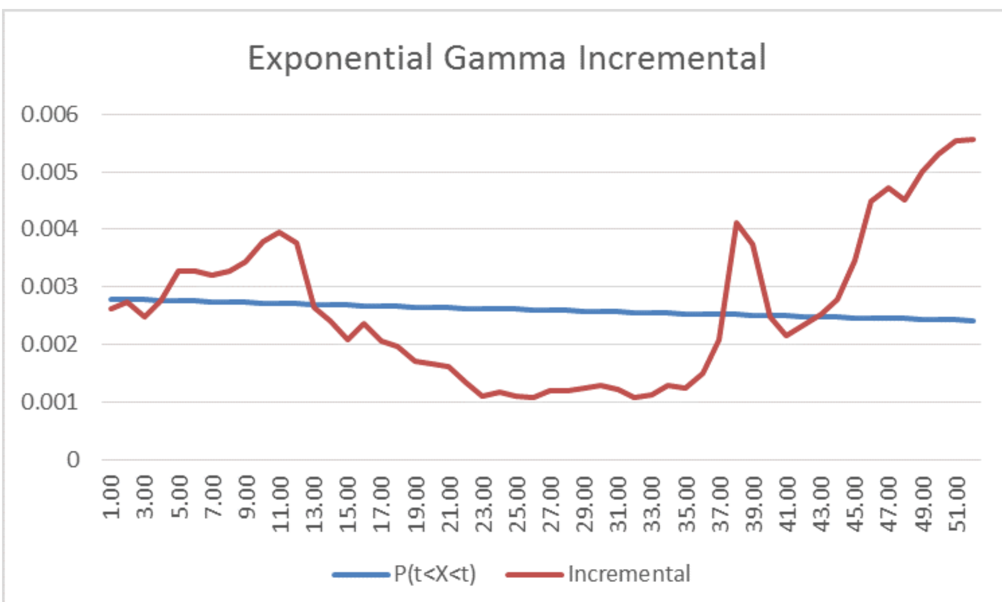

Figure 13: Incremental Tracking Plot

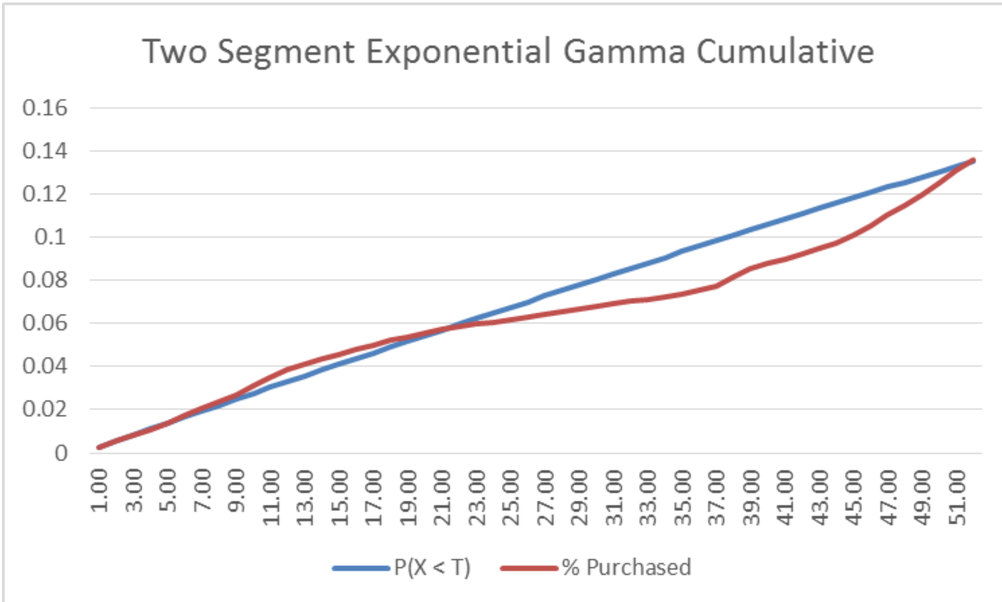

Figure 14: Cumulative Tracking Plot

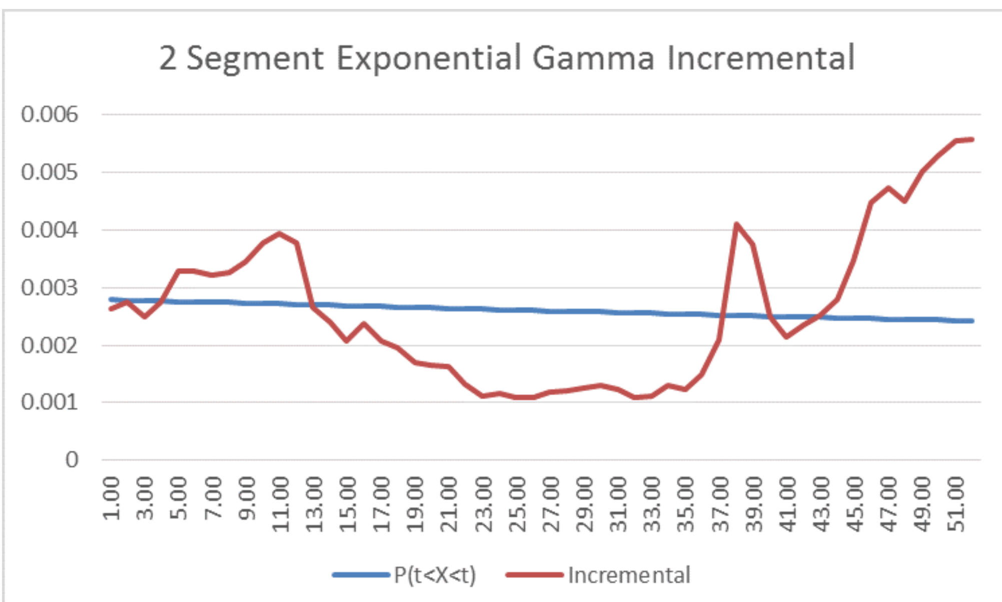

Figure 15: Incremental Tracking Plot

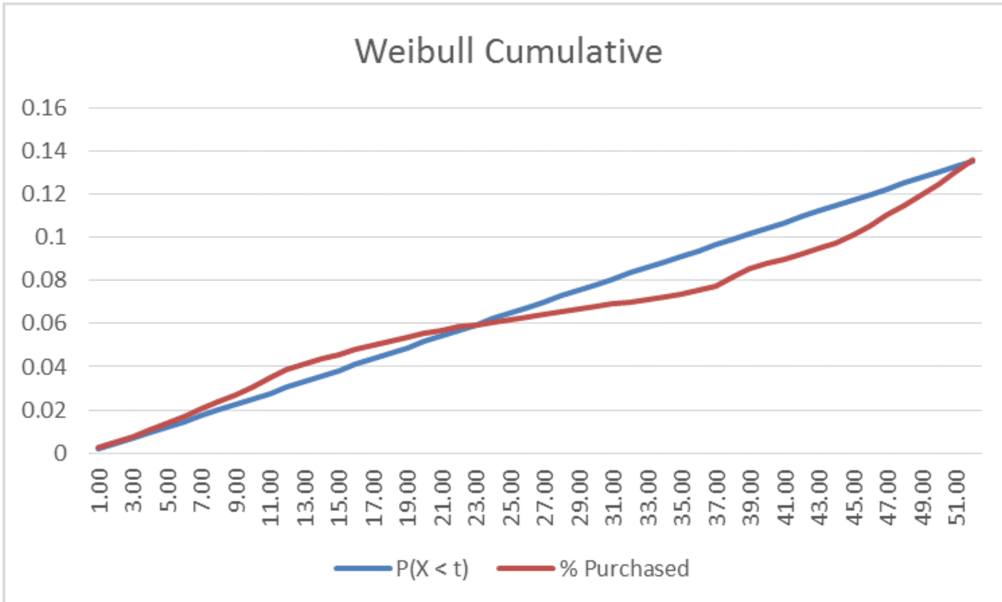

Figure 16: Cumulative Tracking Plot

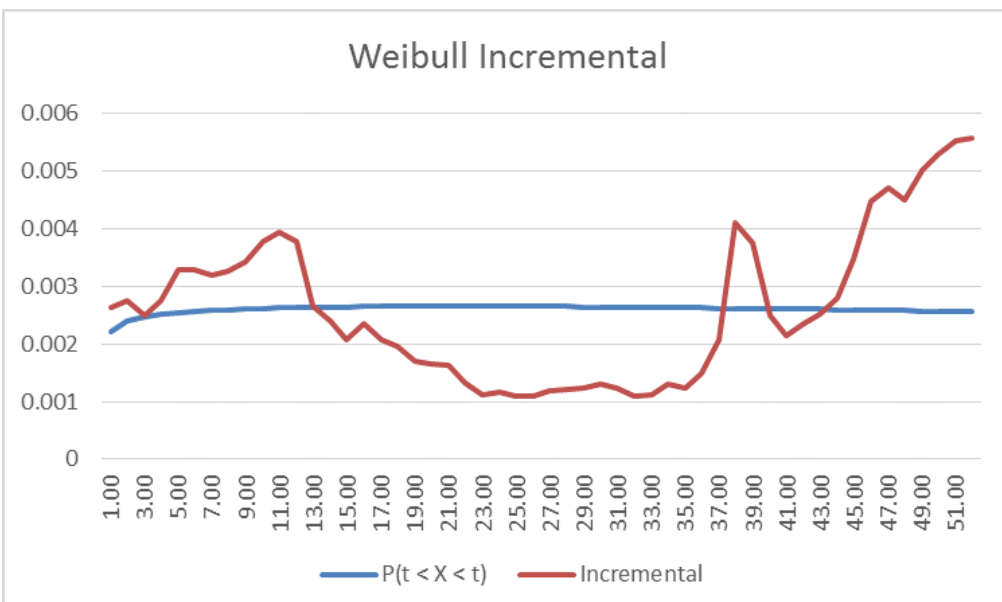

Figure 17: Incremental Tracking Plot

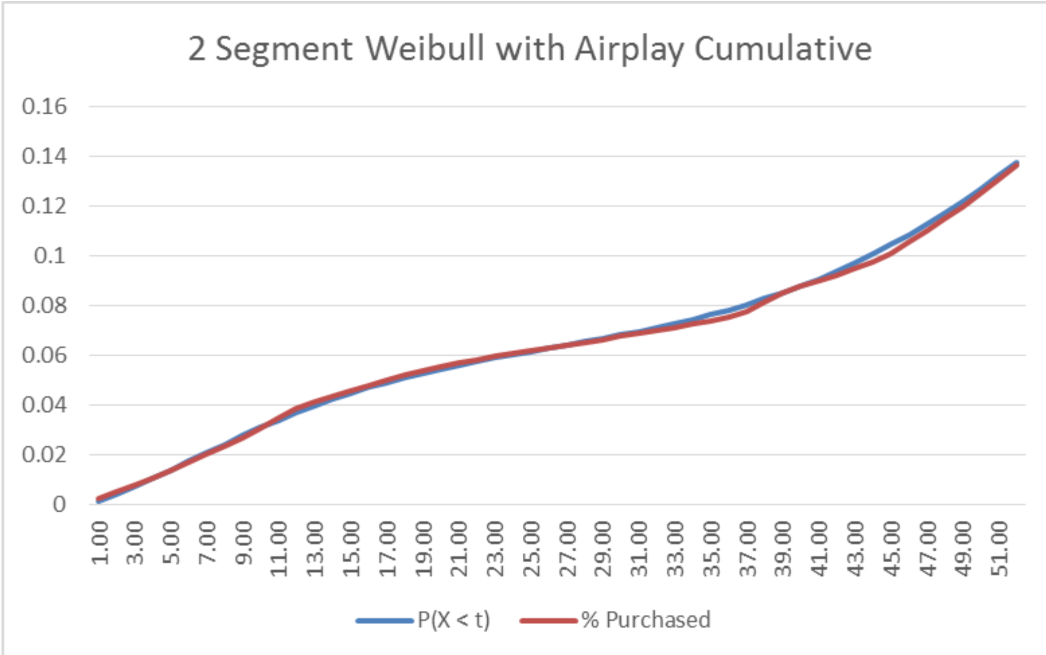

Figure 18: Cumulative Tracking Plot

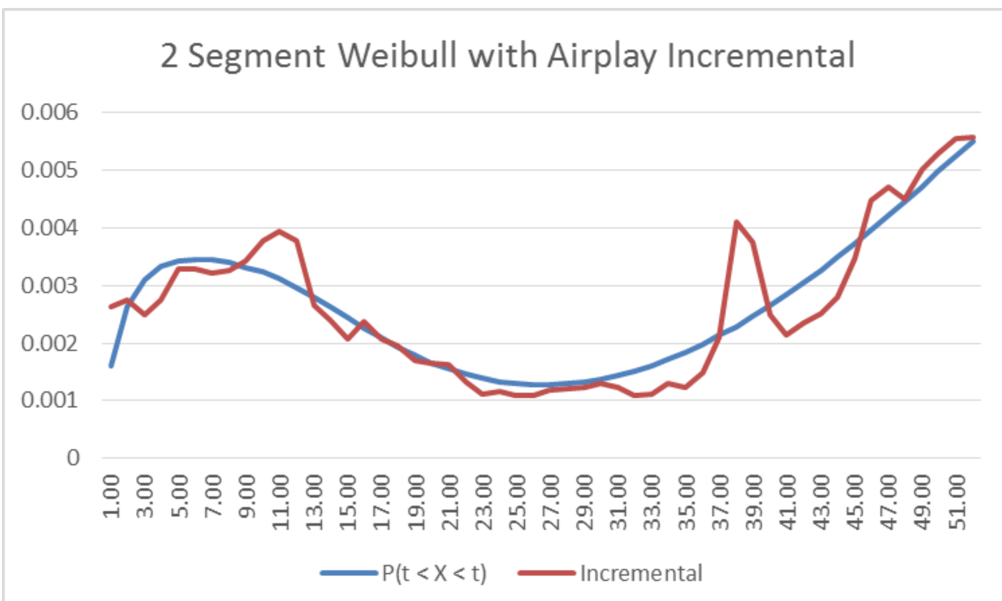

Figure 19: Incremental Tracking Plot

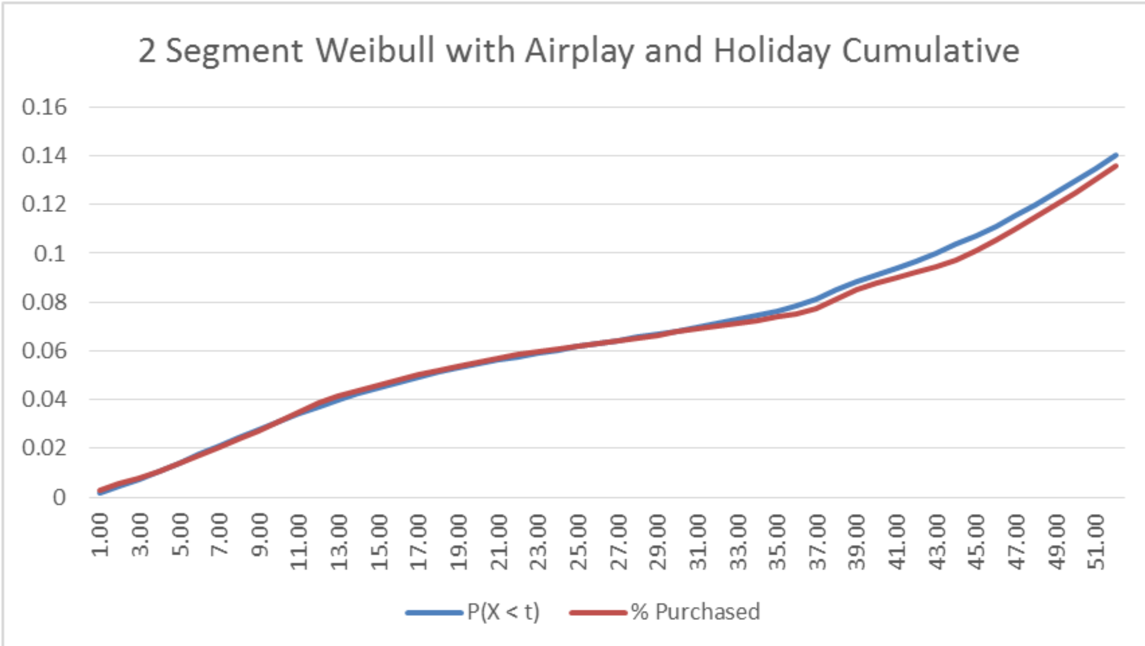

Figure 20: Cumulative Tracking Plot

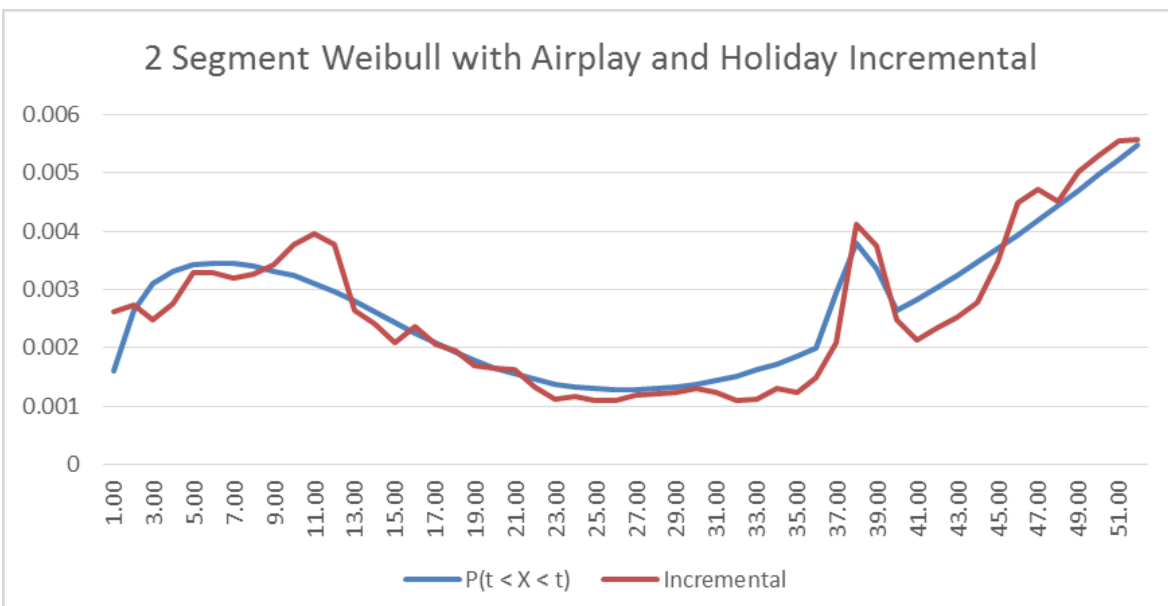

Figure 21: Incremental Tracking Plot

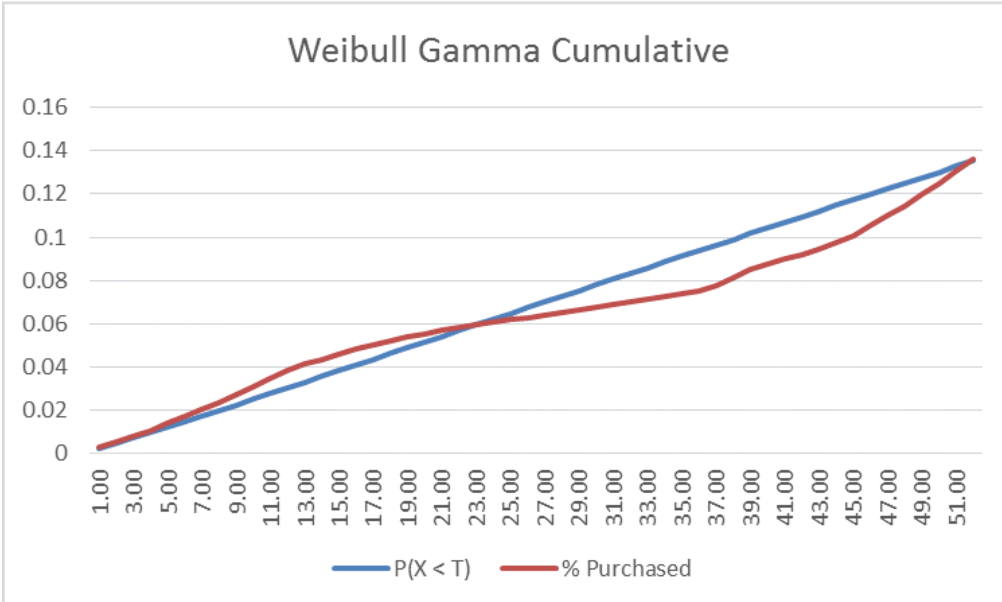

Figure 22: Cumulative Tracking Plot

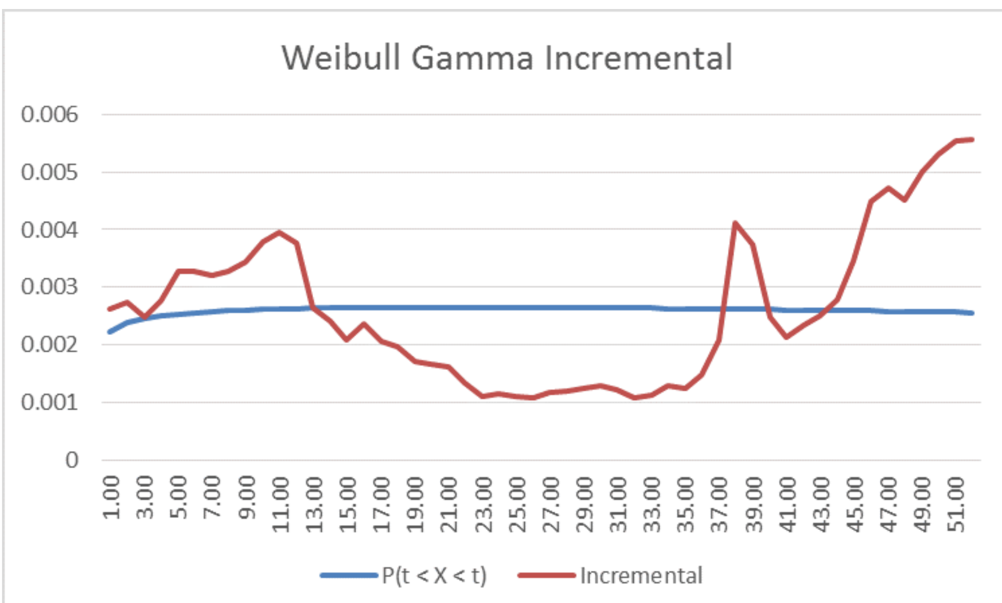

Figure 23: Incremental Tracking Plot

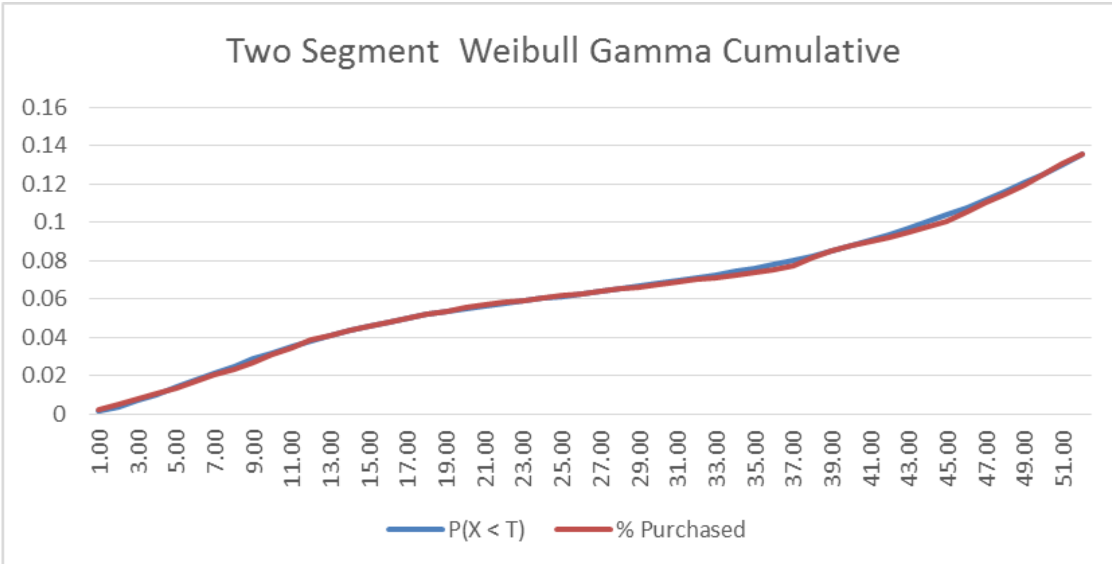

Figure 24: Cumulative Tracking Plot

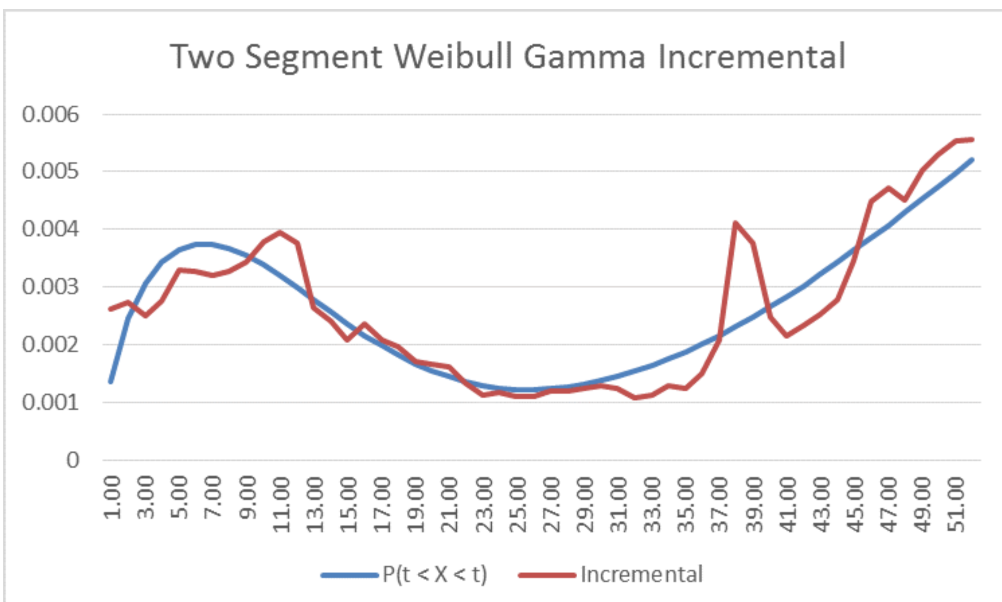

Figure 25: Incremental Tracking Plot

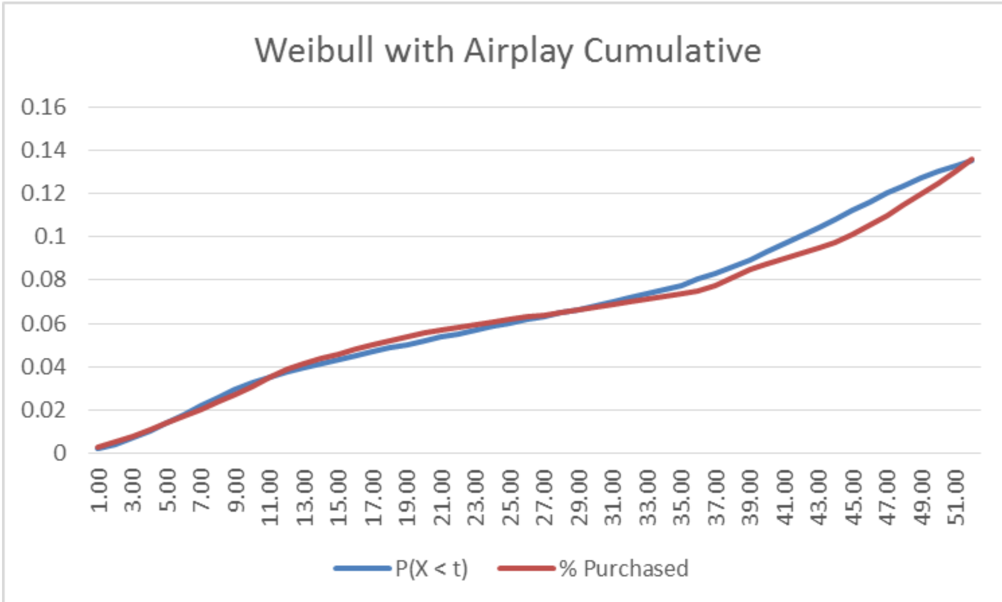

Figure 26: Cumulative Tracking Plot

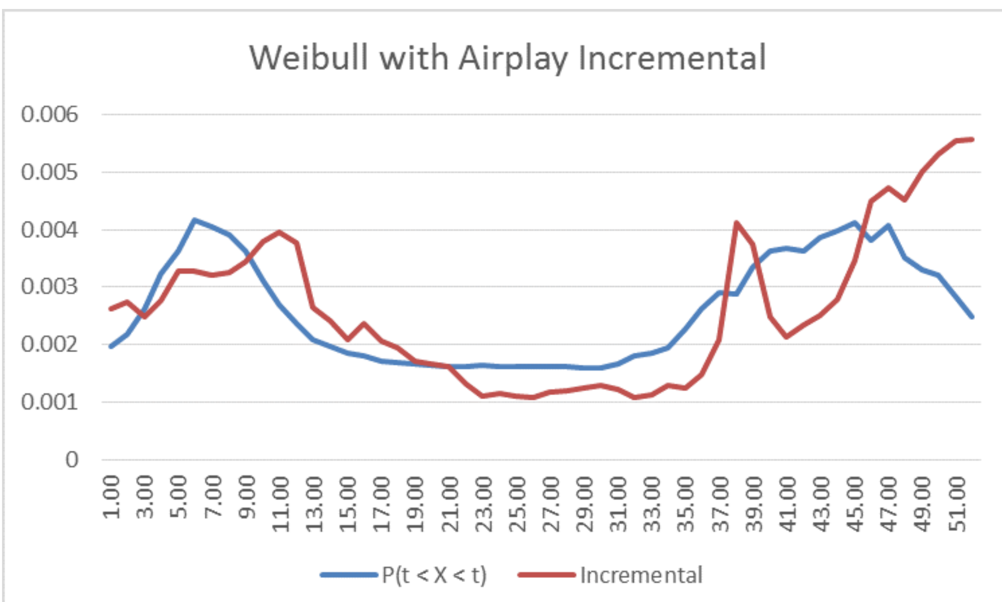

Figure 27: Incremental Tracking Plot