Pricing an Option on Revenue from an Innovation: An Application to Movie Box Office Revenue

TECHNICAL HANDBOOK

Don Chance   Eric Hillebrand   Jim Hilliard

April 23, 2007

Abstract

This manuscript provides technical details about the application of the methods proposed in Chance, Hillebrand, and Hilliard, "Pricing an Option on an Innovation: An Application to Movie Box Office Revenue."

Contents

1 A Detailed Example: Updating "The Others"

1.1 Data Preparation and MLE

The time series of cumulative revenue in Dollars for the first 14 weeks is

$$R = \begin{bmatrix} 14089952 \\ 32168706 \\ 46146680 \\ 59956515 \\ 67516127 \\ 73422887 \\ 80084619 \\ 86694507 \\ 90573083 \\ 93165623 \\ 94543041 \\ 95242858 \\ 96080075 \\ 96471845 \end{bmatrix}.$$

First, we take first differences and divide by \$5, the average ticket price in 2001. This leaves us with the estimated time series of weekly adoption

$$N = \begin{bmatrix} 2817990.4 \\ 3615750.8 \\ 2795594.8 \\ 2761967 \\ 1511922.4 \\ 1181352 \\ 1332346.4 \\ 1321977.6 \\ 775715.2 \\ 518508 \\ 275483.6 \\ 139963.4 \\ 167443.4 \\ 78354 \end{bmatrix}.$$

Next, we scale the series by dividing by its Euclidean vector norm

$$||N|| = 6683042.09.$$

This results in the scaled series

$$X = \begin{bmatrix} 0.421662823994929 \\ 0.541033671793177 \\ 0.418311715367638 \\ 0.413279905070224 \\ 0.226232661702889 \\ 0.176768600933508 \\ 0.19936226381874 \\ 0.19781075481096 \\ 0.116072170383473 \\ 0.0775856253960136 \\ 0.0412212876027858 \\ 0.0209430672652156 \\ 0.0250549671508151 \\ 0.0117243014423678 \end{bmatrix},$$

which we hand to the maximum likelihood routine `movie_gamma.m` with the start value for the parameter vector

$$\theta_0 = (1.5e7/||N||, 0.47, 1.8e5/||N||).$$

This choice of initial values reflects our experiments with the box office data in our sample and may need adjustment for other data. The resulting vector of maximum likelihood parameter estimates is

$$\hat{\theta}_{MLE} = (3.0261, 0.2200, 0.0162).$$

To recover the ML estimates of m, p, and β from these results, we simply multiply the first and the last entry by $||N||$:

$$\hat{m}_{MLE} = \hat{\theta}_{MLE}(1) \times ||N|| = 3.0261... \times 6683042.09 = 20223806.30,$$

which is the estimate of maximum adoption, and

$$\hat{\beta}_{MLE} = \hat{\theta}_{MLE}(3) \times ||N|| = 0.01619... \times 6683042.09 = 108225.58,$$

is the estimate of the volatility parameter. $\hat{p}_{MLE} = 0.22$, since the attenuation parameter is scale-invariant.

1.2 Finding the Initial Estimate

The regression table is copied from the paper:

	log(m)	log(p)	log(β)
Constant	10.56*** (0.71)	0.01 (0.54)	4.97*** (1.12)
Initial Screens	1e-4** (6e-5)	4e-4*** (4e-5)	-5e-4*** (1e-4)
log(Budget)	0.32*** (.04)	-0.14*** (0.03)	0.45*** (0.07)
Rating			
PG	-0.21 (0.20)	0.12 (0.15)	-0.74** (0.32)
PG-13	-0.16 (0.18)	0.33*** (0.14)	-1.08*** (0.29)
R	-0.26* (0.18)	0.43*** (0.14)	-1.11*** (0.28)
SIC	-219.63	-347.53	7.2

The regression is specified in the logs of the parameters m and β. "The Others" was rated PG-13, had a budget of $17 million in production plus $10 million in advertising, and was released on 1,678 initial screens. We obtain the initial estimate thus as

$$\widehat{\log m_0} = 10.56 + 1\text{e-}4 \times 1678 + 0.32 \times \log 27000000 = 16.2034,$$

$$\widehat{\log p_0} = 4\text{e-}4 \times 1678 - 0.14 \times \log(27000000) + 0.33 = -1.3944,$$

$$\widehat{\log \beta_0} = 4.97 - 5\text{e-}4 \times 1678 + 0.45 \times \log(27000000) - 1.08 = 10.7511.$$

This implies an initial estimate of

$$\hat{\theta}_0 = (\hat{m}_0, \hat{p}_0, \hat{\beta}_0) = (10890824, 0.2480, 46681).$$

1.3 The Choice of Priors

The following table is the equivalent of Table 2 in the paper when the scaling by $||N||$ is applied to all series in the sample:

	$\hat{m}$	$\hat{p}$	$\hat{\beta}$
	log-normal	normal	log-normal
μ	0.8339	0.3788	-4.5031
σ	0.2289	0.1436	0.9987

The norm $||N||$ is not known before all revenue observations have arrived, so in practice, it cannot be used in the update. We only use it here for convenience. Any other factor, for example budget/(average ticket price), the adoption necessary to earn the budget, can be used.

We now set the means of these distributions equal to the initial estimates obtained in the last section and keep the variances:

$$m: \quad \log(\hat{m}_0/||N||) = 0.4883$$

$$p: \quad \exp(-1.3944) = 0.2480$$

$$\beta: \quad \log(\hat{\beta}_0/||N||) = -4.9640$$

This results in the following prior distributions that are used in WinBUGS 14. WinBUGS uses the precision $\tau = 1/\sigma^2$ as a measure of the variance.

	$\hat{m}$	$\hat{p}$	$\hat{\beta}$
	log-normal	normal	log-normal
μ	0.4883	0.2480	-4.9640
τ	19.09	48.49	1.0026

1.4 The Update

The model including priors is specified in the WinBUGS file `chhgamma_otr.odc`. The initial values of the Markov chain are calculated as follows, where the subscript $x, 0$ indicates that these are the start values when the series are scaled by $||N||$:

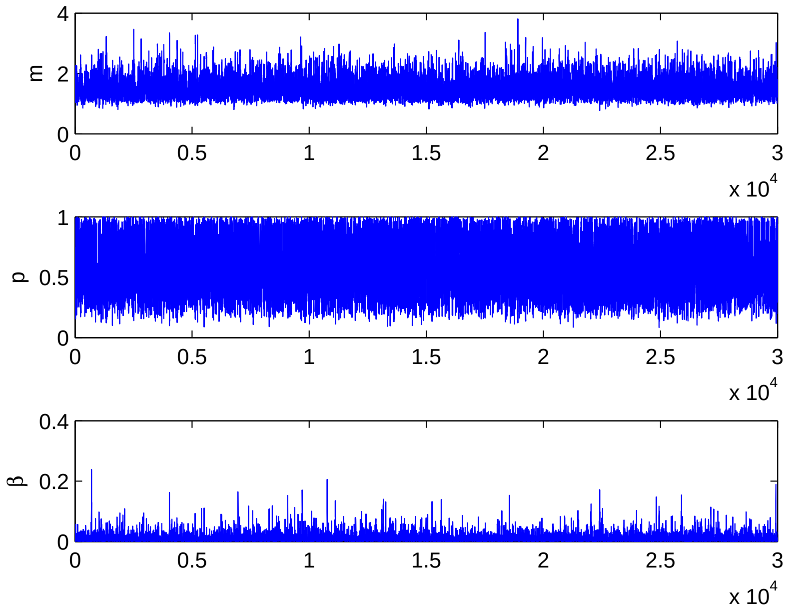
$$m_{x,0} = \hat{m}_0/||N|| = 1.6296$$

$$p_0 = 0.2480$$

$$\beta_{x,0} = \hat{\beta}_0/||N|| = 0.0070.$$

The initial values are specified in the file `chhgamma_otr_seeds.odc`. Finally, the files `otr1.odc`, `otr2.odc`,..., `otr14.odc` contain the time series X from observation 1 through the observation stated in the file name. The `Model` >> `Specification` dialog box is used to load the model from `chhgamma_otr.odc` and then the time series data beginning with `otr1.odc`. After pushing `compile`,

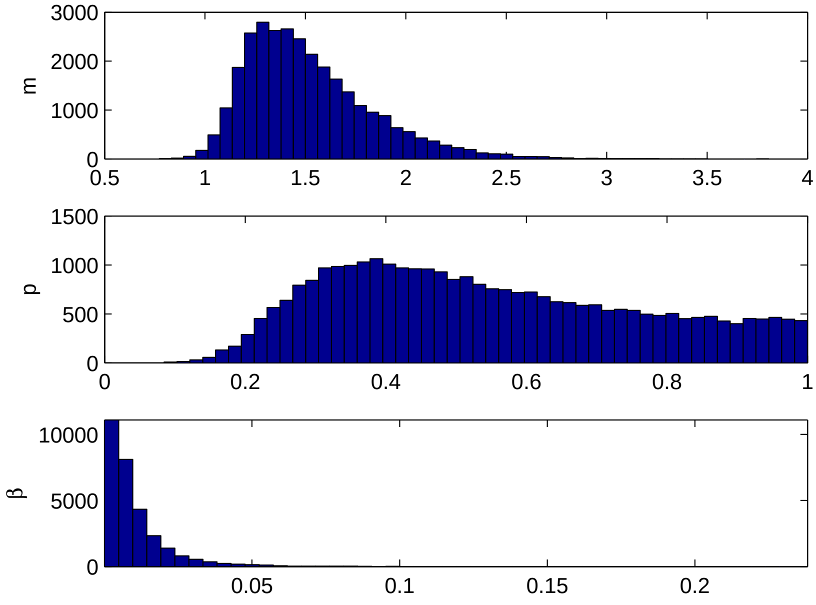
Figure 1: 30,000 values of the Markov chain that samples from the posterior distribution of the update of the initial estimate after the first observation of the revenue series, 14089952, has arrived.



the inits can be loaded from `chhgamma_otr_seeds.odc`. In the `Inference` >> `Sample` dialog box, we set the nodes m, p, and β. In the `Model` >> `Update` dialog box, we specify the number of updates, for this study typically between 10,000 and 30,000. In the `Inference` >> `Sample` dialog box, we can use the `history` or `coda` button to check convergence. These return the series of values for the respective node over the updates of the Markov chain. The following figure shows the 30,000 values of the Markov chain that samples from the posterior distribution of the update of the initial estimate after the first observation of the revenue series, 14089952, has arrived:

The chain has converged to its stationary distribution. The following figure shows the histograms of the samples.

Figure 2: Histograms of the 30,000 values of the Markov chain that samples from the posterior distribution of the update of the initial estimate after the first observation of the revenue series, 14089952, has arrived.



Either from this series or from the `stats` button in the `Inference` >> `Sample` dialog box, we can obtain the mean of the nodes as update of the parameter estimate after the first revenue observation has arrived. In this fashion we continue to run the model for the files `otr2.odc`, `otr3.odc`, ..., `otr14.odc` to obtain the updates as the observations come in. The following table shows the updates in the X-scale that we use in WinBUGS (series normalized by $||N||$) and the translation to the original N-scale that is reported in the paper.

	X-Scale			*Original N-Scale*		
	$\hat{m}$	$\hat{p}$	$\hat{\beta}$	$\hat{m}$	$\hat{p}$	$\hat{\beta}$
$\hat{\theta}_0$	1.630	0.2480	0.0070	10890824	0.2480	46681
week 1	1.519	0.5450	0.0124	10151541	0.5450	82870
week 2	1.935	0.4466	0.0374	12931686	0.4466	249946
week 3	2.254	0.3557	0.0390	15063577	0.3557	260639
week 4	2.572	0.2920	0.0418	17188784	0.2920	279351
week 5	2.801	0.2546	0.0319	18719201	0.2546	213189
week 6	2.943	0.2349	0.0248	19668193	0.2349	165739
week 7	3.143	0.2117	0.0210	21004801	0.2117	140344
week 8	3.359	0.1901	0.0186	22448338	0.1901	124305
week 9	3.402	0.1863	0.0161	22735709	0.1863	107597
week 10	3.365	0.1895	0.0153	22488437	0.1895	102251
week 11	3.274	0.1994	0.0165	21880280	0.1994	110270
week 12	3.175	0.2112	0.0188	21218659	0.2112	125641
week 13	3.160	0.2145	0.0180	21118413	0.2145	120295
week 14	3.129	0.2203	0.0182	20911239	0.2203	121631
MLE	3.026	0.2200	0.0162	20222885	0.2200	108265

2 "The Fellowship of the Ring" and "Ocean's 11"

In the same fashion, we can find the initial estimate for "The Fellowship of the Ring," which had a budget of $149 million, was released on 3359 screens, and rated PG-13.

$$\widehat{\log m_0} = 10.56 + 1\text{e-}4 \times 3359 + 0.32 \times \log 149000000 = 15.8381,$$
$$\widehat{\log p_0} = 4\text{e-}4 \times 3359 - 0.14 \times \log(149000000) + 0.33 = -0.9611,$$
$$\widehat{\log \beta_0} = 4.97 - 5\text{e-}4 \times 3359 + 0.45 \times \log(149000000) - 1.08 = 10.6793.$$

The estimate for the maximum adoption m is very low. Obviously, the regression model is not a very good one to predict the revenue for a blockbuster movie like this. Instead of the regression estimate, we use the adoption that is necessary to earn the budget: $\log(147000000/5) = 17.1965$.

The initial values for the Markov Chain are

$$m_{x,0} = \hat{m}_0/||N|| = 1.1040$$
$$p_0 = 0.3825$$
$$\beta_{x,0} = \hat{\beta}_0/||N|| = 0.0016.$$

The means of the prior distributions are

$$m: \quad \log(\hat{m}_0/||N||) = 0.0990$$
$$p: \quad \exp(-0.9611) = 0.3825$$
$$\beta: \quad \log(\hat{\beta}_0/||N||) = -6.4182.$$

	X-Scale			*Original N-Scale*		
	$\hat{m}$	$\hat{p}$	$\hat{\beta}$	$\hat{m}$	$\hat{p}$	$\hat{\beta}$
$\hat{\theta}_0$	1.104	0.3825	0.0016	29398914	0.3825	42607
week 1	1.552	0.7775	0.0029	41328908	0.7775	77225
week 2	1.677	0.5816	0.0549	44391295	0.5816	1461957
week 3	1.823	0.5210	0.0454	48545490	0.5210	1208977
week 4	1.956	0.4652	0.0370	52087207	0.4652	985290
week 5	2.079	0.4156	0.0318	55362629	0.4156	846817
week 6	2.130	0.4015	0.0277	56720731	0.4015	737636
week 7	2.181	0.3903	0.0243	58078833	0.3903	647096
week 8	2.214	0.3832	0.0218	58957605	0.3832	580522
week 9	2.239	0.3627	0.0213	59623341	0.3627	567207
week 10	2.264	0.3498	0.0201	60289078	0.3498	535252
week 11	2.280	0.3362	0.0194	60715149	0.3362	516611
week 12	2.302	0.3241	0.0187	61300997	0.3241	497971
week 13	2.306	0.3116	0.0186	61407515	0.3116	495308
week 14	2.311	0.2986	0.0187	61540662	0.2986	497971
week 15	2.309	0.2836	0.0194	61487403	0.2836	516611
week 16	2.311	0.2712	0.0196	61540662	0.2712	521937
week 17	2.316	0.2623	0.0193	61673809	0.2623	513948
week 18	2.317	0.2543	0.0191	61700439	0.2543	508623
week 19	2.315	0.2467	0.0190	61647180	0.2467	505960
week 20	2.321	0.2409	0.0185	61806956	0.2409	492645
MLE	2.338	0.2425	0.0202	62259657	0.2425	537915

For "Ocean's 11," which was released on 3075 screens with a reported budget of \$140 million and rated PG-13, we compute the initial estimate as

$$\widehat{\log m_0} = 10.56 + 1\text{e-}4 \times 3075 + 0.32 \times \log 140000000 = 16.8698,$$

$$\widehat{\log p_0} = 4\text{e-}4 \times 3075 - 0.14 \times \log(140000000) + 0.33 = -1.0660,$$

$$\widehat{\log \beta_0} = 4.97 - 5\text{e-}4 \times 3075 + 0.45 \times \log(140000000) - 1.08 = 10.7932.$$

Again, the initial estimate for m is below the necessary adoption for a break-even on the production budget, so we use log(140000000/5) = 17.1477 instead.

The initial values for the Markov Chain are

$$\begin{aligned} m_{x,0} &= \hat{m}_0/||N|| = 1.9455 \\ p_0 &= 0.3444 \\ \beta_{x,0} &= \hat{\beta}_0/||N|| = 0.0034. \end{aligned}$$

The means of the prior distributions are

$$\begin{aligned} m: &\quad \log(\hat{m}_0/||N||) = 0.6655 \\ p: &\quad \exp(-0.9611) = 0.3444 \\ \beta: &\quad \log(\hat{\beta}_0/||N||) = -5.6889. \end{aligned}$$

	X-Scale			*Original N-Scale*		
	$\hat{m}$	$\hat{p}$	$\hat{\beta}$	$\hat{m}$	$\hat{p}$	$\hat{\beta}$
$\hat{\theta}_0$	1.945	0.3444	0.0034	27980841	0.3444	48932
week 1	1.819	0.5675	0.0052	26178633	0.5675	74837
week 2	2.437	0.3608	0.0097	35072749	0.3608	139600
week 3	3.009	0.2617	0.0087	43304843	0.2617	125208
week 4	3.223	0.2461	0.0231	46384682	0.2461	332450
week 5	3.165	0.2437	0.0206	45549959	0.2437	296471
week 6	3.097	0.2586	0.0194	44571319	0.2586	279200
week 7	3.095	0.2472	0.0174	44542535	0.2472	250417
week 8	2.996	0.2634	0.0183	43117750	0.2634	263369
week 9	2.940	0.2750	0.0184	42311811	0.2750	264809
week 10	2.902	0.2859	0.0187	41764923	0.2859	269126
week 11	2.873	0.2962	0.0190	41347562	0.2962	273444
week 12	2.842	0.3069	0.0202	40901417	0.3069	290714
week 13	2.833	0.3125	0.0200	40771891	0.3125	287835
week 14	2.831	0.3150	0.0194	40743107	0.3150	279200
week 15	2.834	0.3138	0.0183	40786283	0.3138	263369
MLE	2.557	0.3144	0.0186	36796883	0.3144	267687

3 Imitation Effect and $q \neq 0$ (Section 6.2)

The literature has documented that there is an imitation effect in adoption for some movies, which translates to $q \neq 0$ in the Bass model. This imitation effect is particularly pronounced in platform releases that rely more on word-of-mouth than on large marketing efforts. In our study, we focus on wide releases with larger marketing budgets in which the imitation effect is not so prevalent, and assume $q = 0$ in the Bass model. The reason for this procedure is that there are well documented econometric problems with the fully specified Bass model. Boswijk and Franses (2004) found that if one adds a Brownian Motion error to a deterministic Bass mean function, one cannot estimate q consistently if there are no data beyond the local maximum available. Even though our model has a different stochastic structure, we expect a similar result to hold, which will adversely affect the update procedure, where only few observations are available. In the case of movie revenues, we also have to deal with substantial small-sample effects. We illustrate this problem in a few simulations.

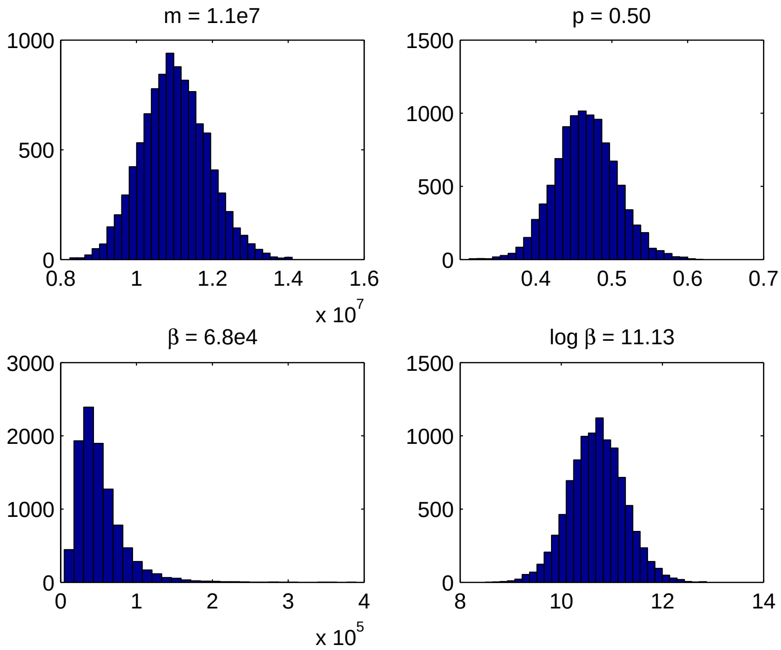
Figure 3 reports the histograms of parameter estimates from 10,000 replications of time series of length $T = 20$ from the model when we set $q = 0$. The data-generating parameter vector is $\theta = (1.1e7, 0.5, 0, 6.8e4)$. As can be seen from the histograms, small-sample distortions seem reasonable: there is a small downward bias in p and β and the distributions seem to be normal.

Contrast this with the estimator histograms in Figure 4, where we use the same simulation and estimation setup, only that we set $q = 0.90$. The quality of the estimators is much worse. The distributions of m, p, and β exhibit bimodality and q has a mode on the wrong end of the unit interval.

These strong small-sample distortions result in a bad fit of the option price to the discounted payoffs from the options as can be seen in Figure 5 (paralleling Figure 3 in the paper). This leads us to drop the q-term from the model for the movie revenue application.

As is well-known in statistical modeling of markets, the assumptions need not necessarily be palatable to all observers of the markets. The common assumption in finance that everyone borrows and lends at the same risk-free rate is hardly correct but widely assumed and generally considered innocuous in most models. In the end, what matters is model fit and in our example of movie revenues, the model fit (Figure 3 in the paper) is very good. If the fit is good,

Figure 3: Histograms of Parameter Estimates from 10,000 Simulations of a Gamma Process with Length $T = 20$ and Parameter Vector $\theta = (1.1e7, 0.5, 0, 6.8e4)$.



either the assumption is insignificant or its effects are second- or higher-order and dominated by more powerful first-order effects.

Having said that, however, we do not mean to imply that q should be dropped in general from the innovation model. This is why we chose the Bass model and not just an exponential function for the time-varying mean. In applications where the sample sizes are not so heavily restricted as in the movie example, the model can be estimated with reasonable accuracy. To demonstrate this point, we set up a simulation where the length of the sample paths is $T = 200$. For the data-generating parameter vector, we choose $\theta = (1e8, 5\text{e-}5, 0.08, 5\text{e}4)$. A typical trajectory of the increment process is plotted in Figure 6.

In this case the small sample distortions are not severe. Figure 7 shows the

Figure 4: Histograms of Parameter Estimates from 10,000 Simulations of a Gamma Process with Length $T = 20$ and Parameter Vector $\theta = (1.1e7, 0.5, 0.9, 6.8e4)$.

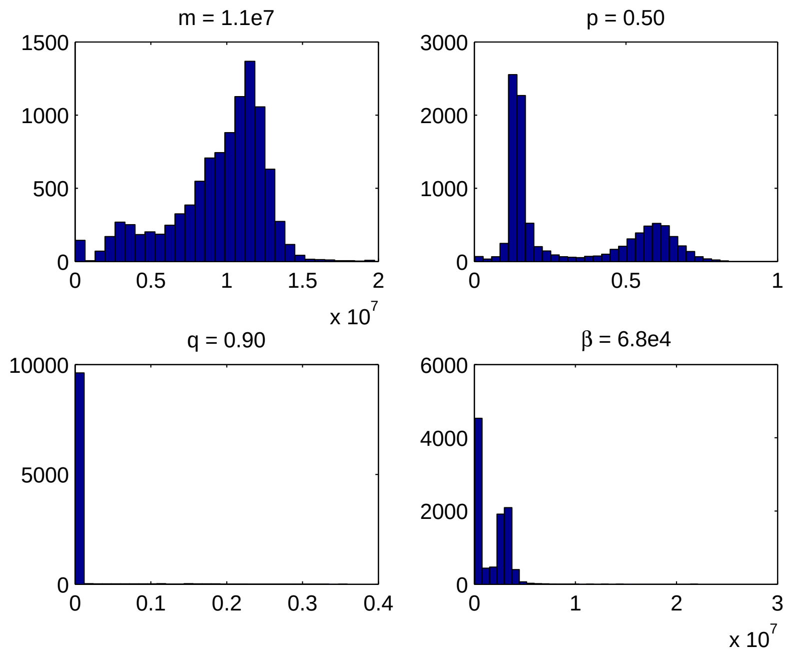


Figure 5: Model Prices versus Average Discounted Payoffs.

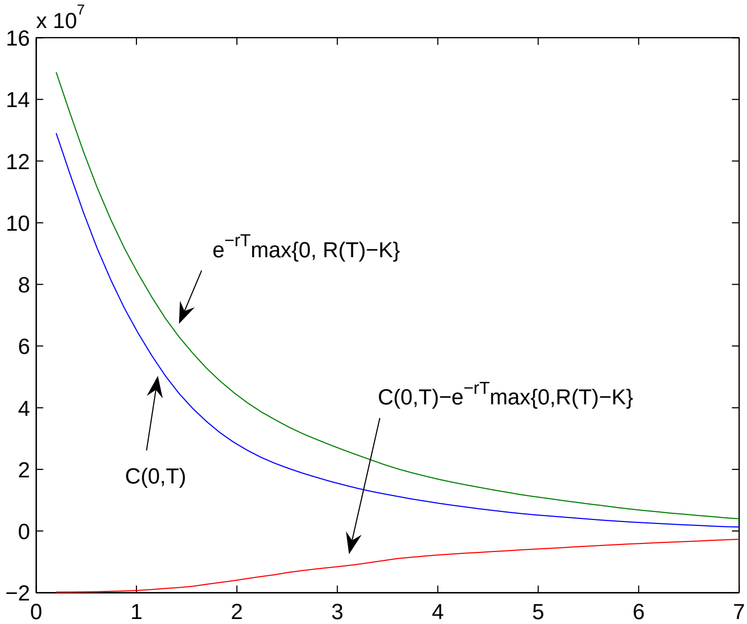


Figure 6: Typical Trajectory from the Gamma Process with Length $T = 200$ and Parameter Vector $\theta = (1e8, 5e-5, 0.08, 5e4)$.

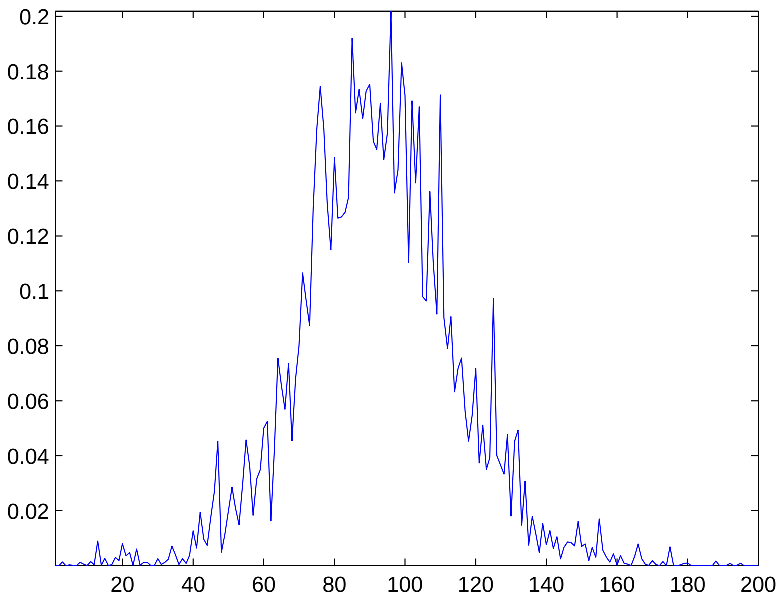
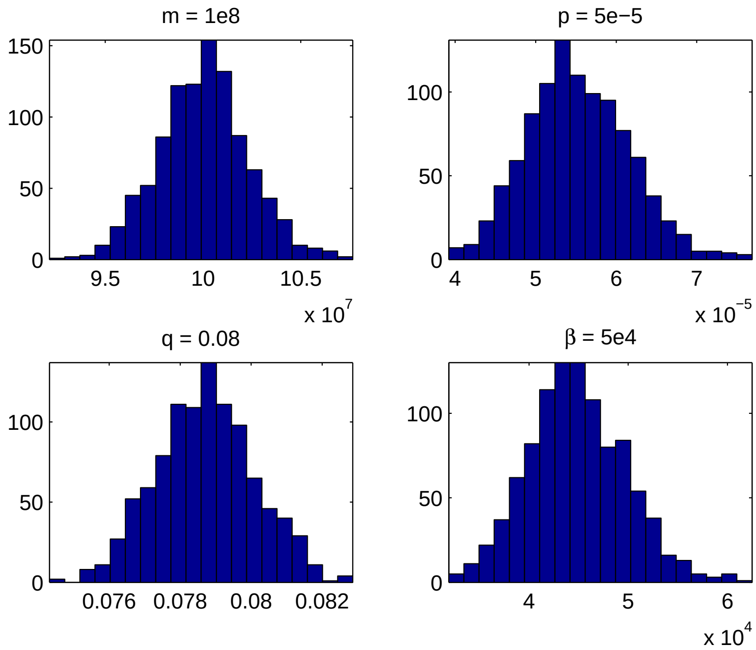


Figure 7: Histograms of Parameter Estimates from 1,000 Simulations of the Gamma Process with Length $T = 200$ and Parameter Vector $\theta = (1e8, 5e-5, 0.08, 5e4)$.

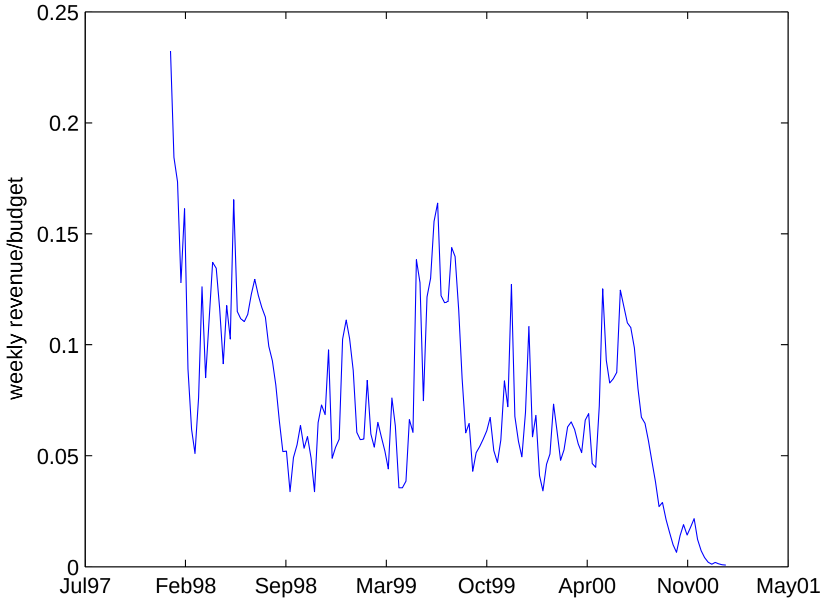


histograms of the parameter estimates from 1,000 simulations of the process. Small-sample biases in p, q, and β are present but not very pronounced. All distributions are unimodal and seem to be normal.

4 Independence of Movie Returns from Stock Market Returns (Section 6.1)

We divided the interval 01-Jan-1998 through 25-Jan-2001 into 161 weekly bins and recorded for each bin the revenue divided by budget generated by the 244 movies in our sample. Figure 8 shows the resulting movie returns index time series.

Figure 8: Weekly movie returns (revenue/reported budget) index based on 244 movies from 01-Jan-1998 through 25-Jan-2001.



We clipped the first 15 and last 15 observations, which show pronounced small sample bias. Then, we used weekly returns on the S&P500 and ran bidirectional Granger causality tests as well as regressions including contemporaneous values. There was no correlation either way at all reasonable lags (up to 30).