Random Forest for Digital Bermudian Option Pricing
Machine Learning Term Project, Spring 2019

Anthony Maylath; aem578

May 15, 2019

1 Appendix A: Sensitivities to Random Forest Parameters

Tables 1 and 2 show the mean absolute regression errors with varying number of trees and minimum samples per leaf respectively. The number of trees does not have too much of an effect while the performance seems to improve slightly with a relatively small minimum number of samples per leaf. The baseline parameters can be seen in table 4 in Appendix B.

Num Trees	Wedding Cake	Double Digital
1	0.289738	0.390846
5	0.28191	0.386902
10	0.281548	0.38737
25	0.291477	0.39281
50	0.286104	0.393945
100	0.28583	0.393417
250	0.28381	0.389033
500	0.289321	0.39503

Table 1: Absolute Mean Error - Varying Number of Trees

Min Samples Per Leaf	Wedding Cake	Double Digital
0.25	0.288254	0.468675
0.1	0.282377	0.397108
0.05	0.289199	0.392853
0.01	0.289994	0.385432
0.001	0.281837	0.395174
0.0001	0.285335	0.394153
≈ 0	0.286628	0.390055

Table 2: Absolute Mean Error - Varying Number Min Samples per Leaf (%age)

2 Appendix B: Simulation Parameters

For the wedding cake, the exact baseline was 0.15060399666443355. For the double digital, the exact baseline was 0.5558366192710213. The wedding cake payoff gives 1$ is the stock is within 1 of the strike and 0.5$ if the stock is within 5 of the strike. The two strikes of the double digital are 90 and 100. I used 250 samples (average of 250 simulations) for the wedding cake and 600 samples for the double digital. The settings for tables 5 and 4 only uses one simulation with many paths. **Note that I re-number the tables found in the main document in Appendix C.**.

Paths	Continuation Paths	Initial Stock Price	Strike	Volatility	Interest Rate
2,000,000	1,000,000	105.5	100	15%	3%

Table 3: Baseline Convergence Parameters

Payoff	Number of Trees	Max Depth	Minimum Samples per Node
Wedding Cake	50	5	1%
Double Digital	50	3	1%

Table 4: Random Forest Settings for Figure 9

Paths	Initial Stock Price	Strike (Close/Far)	Double Digital Strike	Volatility	Interest Rate
1,000,000*	94.5	100/120	90	15%	3%

Table 5: Parameters for Tables 7, 8 and 9

Paths	Continuation Paths	Initial Stock Price	Strike	Double Digital Strike	Volatility	Interest Rate
2,000,000	1,000,000	95	120	75	15%	3%

Table 6: Parameters for Table 10

*Due to a high computational burden, I only use 1 million paths for table 8 while tables 7 and 9 use 200,000 paths. Table 8 gets priority as it contains the most critical results.

3 Appendix C: Recap of Tables and Figures

LSM Step	Wedding Baseline	Wedding RF	WC Improvement	DD Baseline	DD RF	DD Improvement
1	0.265676	0.264408	0.001268	0.539953	0.536331	0.003622
2	0.285662	0.286189	-0.000527	0.477726	0.483434	-0.005708
3	0.299417	0.297026	0.002391	0.431221	0.431035	0.000186
Avg	0.283585	0.282541	0.001044	0.48297	0.4836	-0.0006

Table 7: Regression Errors at Each Exercise Date

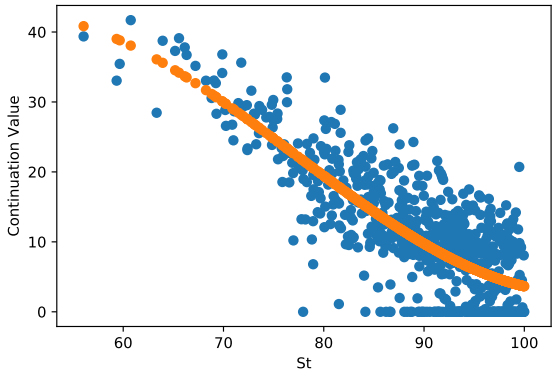

Figure 1: Regression for Vanilla Put

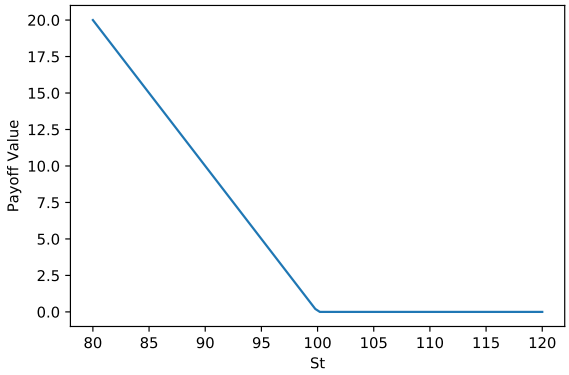

Figure 2: American Put Payoff Function

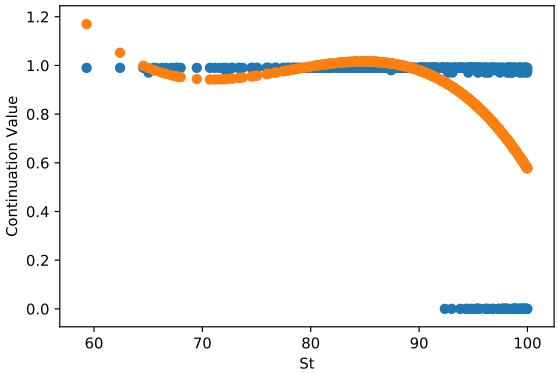

Figure 3: Regression for Digital Put

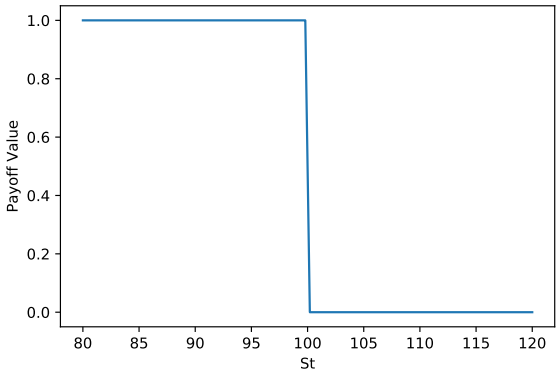

Figure 4: American Digital Put Payoff

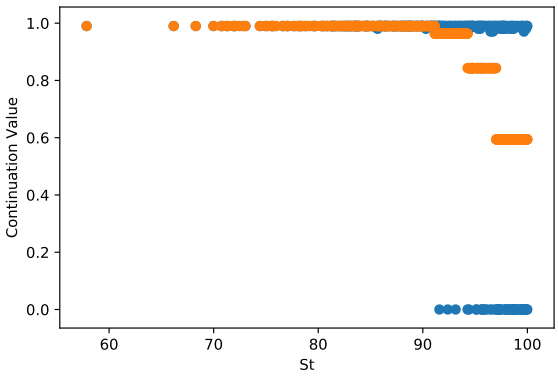

Figure 5: Decistion Tree for Digital Put

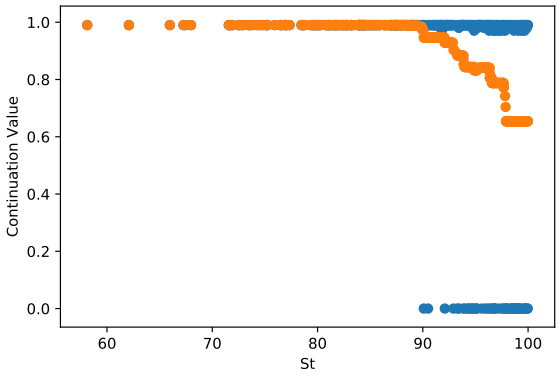

Figure 6: Random Forest Prediction for Digital Put

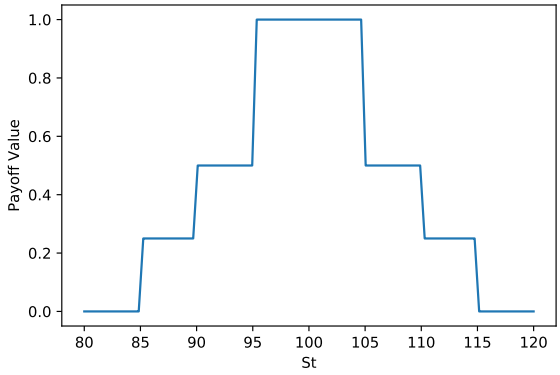

Figure 7: Wedding Cake Payoff

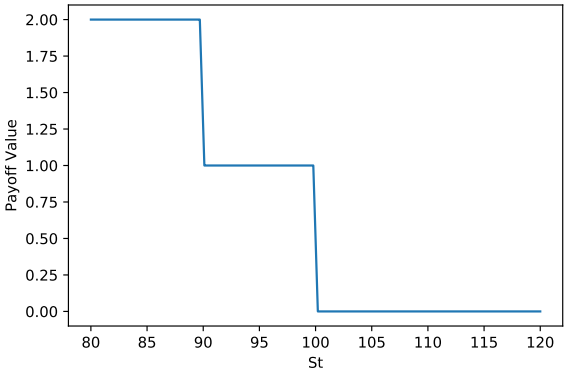

Figure 8: Double Digital Payoff

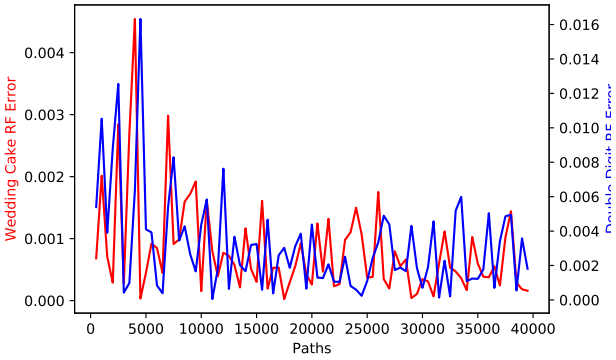

Figure 9: LSMRF Convergence for Wedding Cake and Double Digit

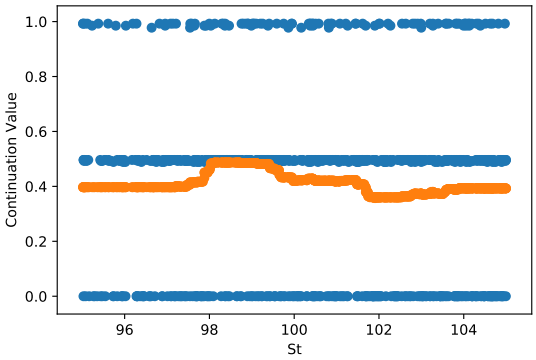

Figure 10: Wedding Cake Regression Fit

TimeSteps	Baseline WC	Random Forest WC	Improvement WC	Baseline DD	Random Forest DD	Improvement DD
4	0.099282	0.098192	0.00109	0.324702	0.319951	0.004751
12	0.07695	0.073791	0.003159	0.309063	0.288816	0.020247
52	0.085752	0.082098	0.003654	0.274474	0.230078	0.044396

Table 8: Regression Errors for Varying Number of Exercise Dates - Far Strikes

TimeSteps	Baseline WC	Random Forest WC	Improvement WC	Baseline DD	Random Forest DD	Improvement DD
4	0.282667	0.282136	0.000531	0.482253	0.484866	-0.002613
12	0.273115	0.276103	-0.002988	0.374536	0.364623	0.009913
52	0.304701	0.302849	0.001852	0.304125	0.297083	0.007042

Table 9: Regression Errors for Varying Number of Exercise Dates - Close Strikes

Error	Baseline WC	Random Forest WC	WC Improvement	Baseline DD	Random Forest DD	DD Improvement
Test	0.040996	0.04016	0.000836	0.100562	0.091774	0.008788
Training	0.052434	0.051018	0.001416	0.118959	0.104728	0.014231

Table 10: Mean Absolute Errors Against the Monte Carlo Tree Method