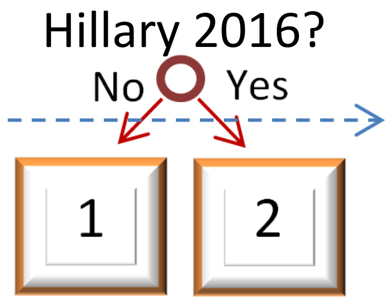
Combinatorial Binary Prediction Markets

Paul Sztorc
truthcoin@gmail.com
Version 1.1

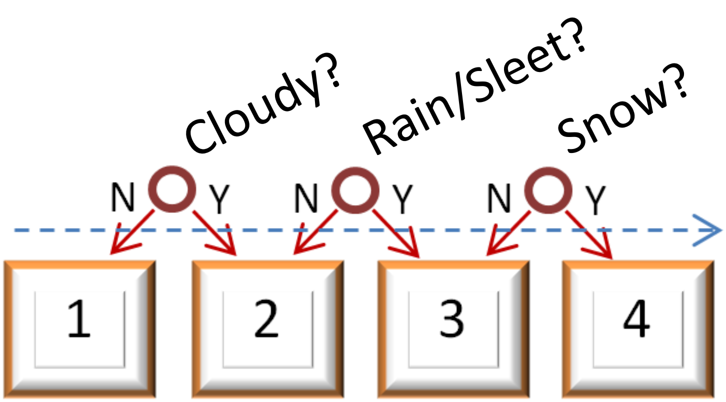
Summary

This document is a short guide to combinatorial use of binary prediction markets (PMs). While binary PMs are powerful on their own (capable of accurately estimating the probability of any defined event), when *combined* they yield even greater insights. PMs combined within-dimension can assess the probabilities of events with any number of mutually exclusive states. PMs combined across-dimension can assess joint and marginal probabilities of multiple variables. “Chained” PMs can elicit the probability density function (higher moments, multi-modality, hazard function for time series, etc.) of an event.



Example 1 – Canonical Binary (K=1, N=2, D=1)

Consisting of a single Yes/No decision requirement, these PMs were popularized by InTrade[1] . Participants who honestly doubt the event will happen can profit by purchasing a share of State 1, whereas those who find it likely can purchase a share of State 2. Assuming market efficiency, the price of State 2 represents the likelihood of the actual event. Above, the features of this contract are represented graphically: two States (yellow squares), one required Decision (red circle) and one Dimension (blue dashed arrow).



Example 2 – Categorical (K=k, N=(k+1), D=1)

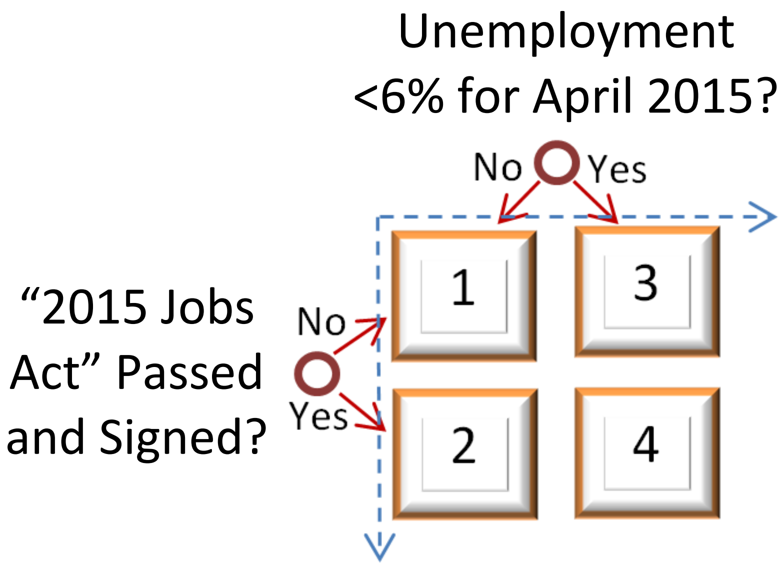
A variable of interest might take on more than just two mutually-exclusive States. This indicates multiple categories within a variable (one of [A, B, or C]), but not gradations across a continuous variable (which of [>2, >3, and/or >4]).

For example, the weather[2] could be described as Clear/Sunny, Cloudy/Overcast, Rain/Sleet/Hail, or Snow. Using three Decisions “Was it Cloudy/Overcast on June 21st, 2015?”, “Was it Raining (including Sleeting, Hailing) on June 21st, 2015?”, “Was it Snowing (but not Sleeting or Hailing) on June 21st, 2015?”, we could partition a PM into the 4 States desired. Clear/Sunny would be the null state (‘1’).

---

[1] http://www.youtube.com/watch?v=N_DWqeR9jqc

[2] Can Weather Forecasts Be Improved? (wired.com)



Example 3 – Multidimensional Binary (K=k, N=2^k, D=k)

Contracts that span two Dimensions (blue, dashed arrows) have a remarkable property: the ability to reveal the relationship between the two dimensions. For example, congressmen of 2014 can estimate the efficacy of a Jobs Bill they are considering for 2015, by constructing this Market and observing the relationship between the passage of the bill (dimension 1) and the future employment rate (dimension 2). If the relationship is weak, the bill will be ineffectual.

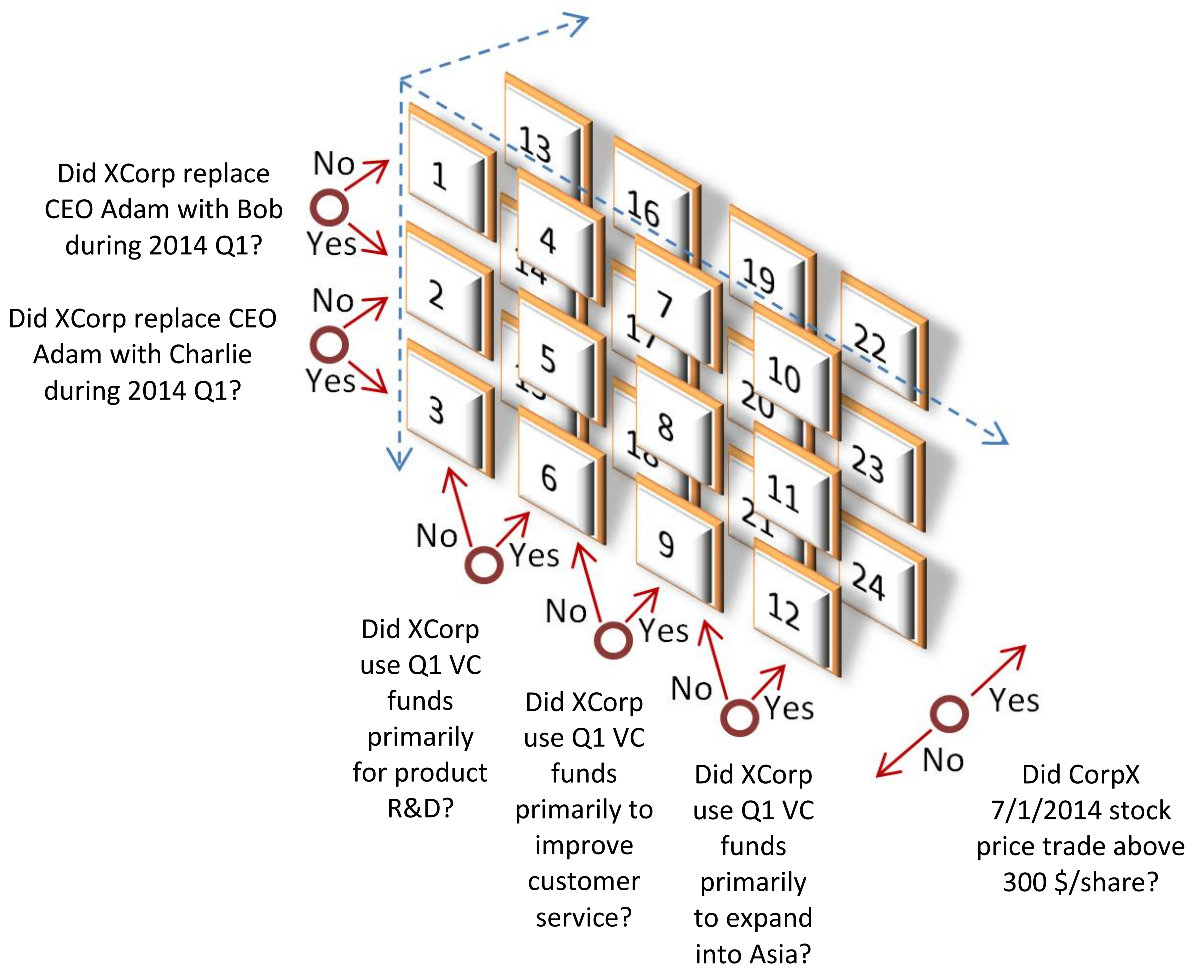
	No	Yes	
No	0.35	0.05	0.40
Yes	0.15	0.45	0.60
	0.50	0.50	1.00

	No	Yes	
No	0.15	0.45	0.60
Yes	0.35	0.05	0.40
	0.50	0.50	1.00

	No	Yes	
No	0.25	0.25	0.50
Yes	0.25	0.25	0.50
	0.50	0.50	1.00

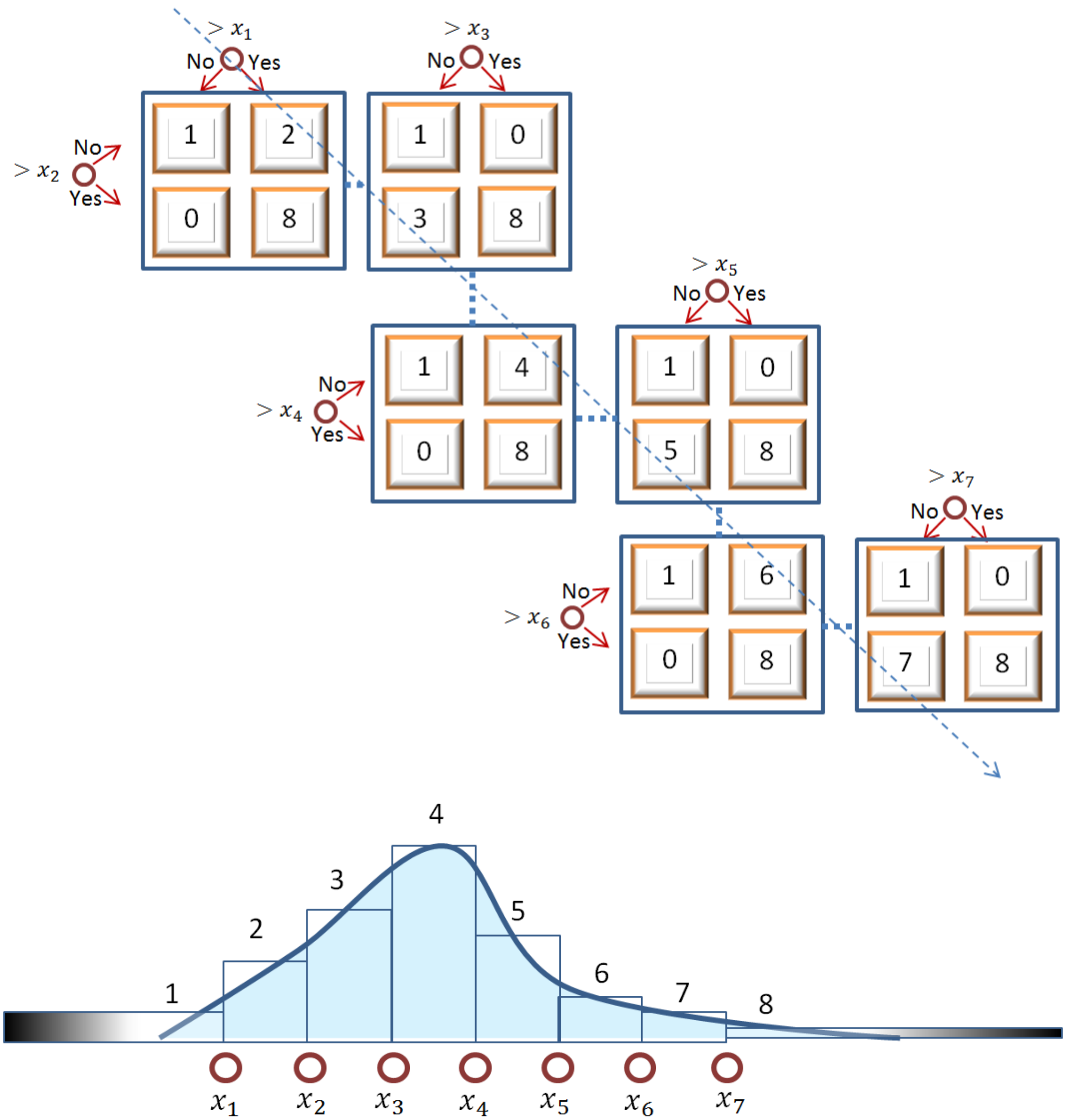
Relationships Between Two-Dimensional Contracts

A layout of three 2-d PMs, with the Decision labels ("No"/"Yes", top and left axes), joint-probabilities (center 4 squares), total probabilities (sums, right and bottom axes), shaded by density. Left: probability is relatively higher in the [No,No] cell and [Yes,Yes] cell, indicating a positive relationship between the variables. Center: when probability for one variable is high in Yes, probability in the second variable is high for No (and vice-versa), indicating a negative relationship. Right: no accumulation of excess probability anywhere, indicating that the variables have no relationship to each other.



Example 4 – Multidimensional Categorical (K=k, N=$\prod(k_d+1)$, D=d)

This example illustrates the flexibility of market creation opportunities, which can take on any number of Decisions, States, and Dimensions. Corporation X "CorpX" could use this device to cheaply assess which of their CEO-replacement or Fund-use decisions would maximize the probability of achieving a price target for their traded shares. Although Traders have many ways of losing, the risk is compensated by reward: early-entrants can multiply their investment by N (24 in this case). Traders are free to ignore Dimensions: By purchasing states 1-12 and 13-24 (for example) in groups, a Trader can experience the same cost, risk, and reward as a binary market on the Decision "Did CorpX 7/1/2014 stock price trade above 300 $/share?".

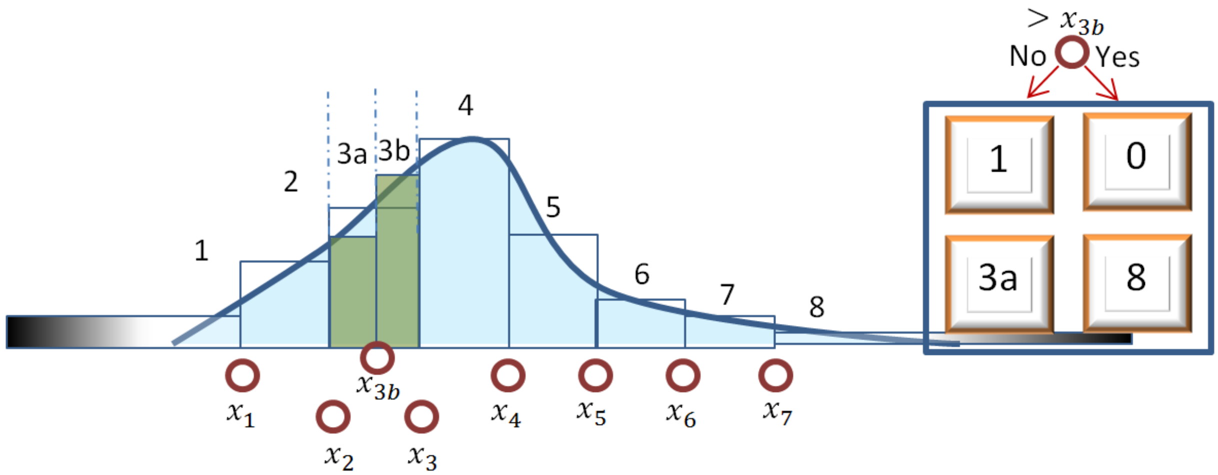


Example 5a – Chained Contracts (K=k, N=4(k-1), D=1)

This is an inefficient form of 'chained contracts' that I provide for instructive value (the non-redundant form is Example 6). To estimate the continuous value of a single variable, multiple Markets must be funded.[3] Note regions 1 through 8 correspond to the prices of states 1 through 8, forming a probability density function. State 0 indicates that the state is logically impossible and should always trade at a price of zero.

"Will the BTC/USD exchange rate be above (500, 1000, 1500, 2000, 2500, 3000, 3500)?"

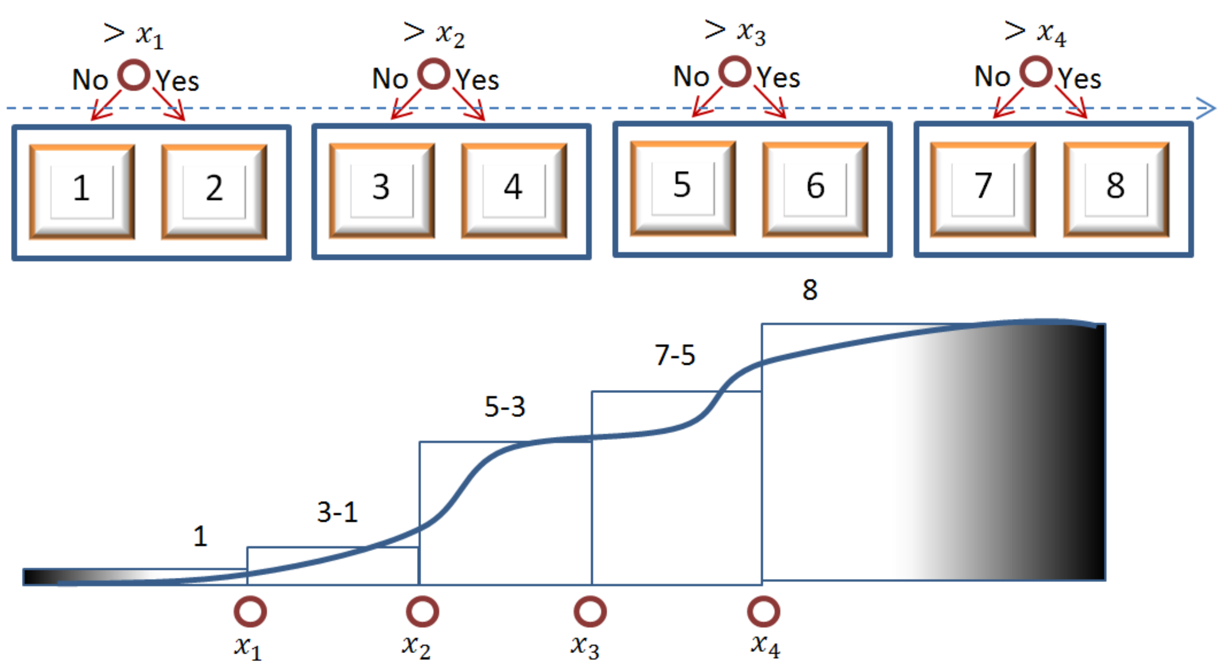
[3] This is a requirement because there will be at least one winner (either yes or no) for each range, and all winners will need to be paid the unit price.



Example 5b – Adding A Link

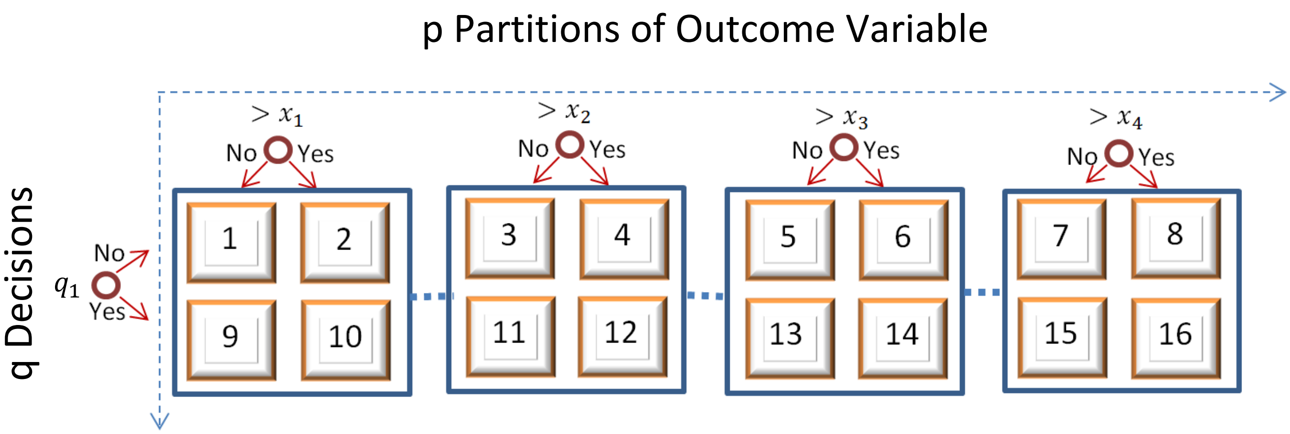
Again, this continues the example which is simply easier to visually understand, and is inefficient compared to 6. Here I describe an Author, noticing high trade volume, adding a new market to further clarify the density function. The price of 3a is given directly by the new market, and the price of 3b is implied: p(3b) = p(3) – p(3a).

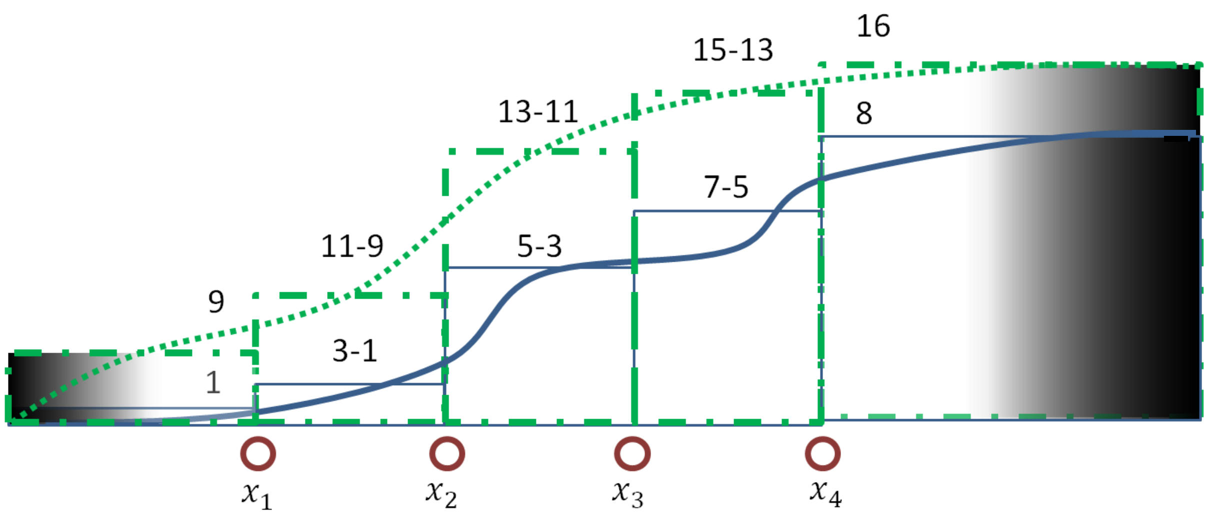
"Will the BTC/USD exchange rate be above (1250)?"



Example 6 – Efficient Chained Contracts (K=k, N=2k, D=1)

This is the efficient example which, without Example 5, may have been harder to grasp visually. A group of 2-State Markets define the cumulative distribution function of a single variable. Prices of the various regions require arithmetic to compute, making them less intuitive.





Example 7 – Complex Advice (K=(p+q), N=(2p)(2^q), D=(1+q))

This set of chained Markets is highly complex yet highly informative. One possible example: (on the left) a Yes / No policy decision, such as "Will the United States declare war against the Central Powers before Jan 1st, 1918?", and (across the top), a possible continuous outcome "Did more than [10,000; 100,000; 1,000,000; 10,000,000] Americans die in 1917?". This market will clarify the relationship between the decision and the outcome by indicating distributional features such as variance/uncertainty, best/worst-case-scenario, and bimodality. This technique, of course, can take on an unlimited number of decision-dimensions to examine the interaction of multiple separate decisions.

Note that a market powered by LMSR (as in Truthcoin markets) allows one to selectively trade only on the information they possess (marginal, joint, total), and also prevents uninformed traders from throwing their money away: information is aggregated producing prices that make it impossible for Traders to spend anything on impossible States.