April 15, 2015

Linear model by ordinary least squares. Logistic Regression model for forecasting the binary outcome (but glm also has a lot of other features.)

Decision Trees

- They are useful abstractions for the decision making process. First step to thinking systematically about decisions.

Functions of Random Variables

- Notation:
  - X: random variable
  - Little x: Some possible outcome
  - Omega: Set of all possible outcomes
- Probability distribution: Distribution is complete description of uncertainty. This is contained in the table (brute force list), but was very complex. Function helps us fill in table for any row, but also limited because it sticks you with certain assumptions. Probability mass functions always involved simplifying assumptions. Quoting whole distribution is often difficult. This is where idea of moments comes in.
- **Moment**: Some simplified description of a probability distribution. Designed to pick out some feature of distribution. (ex: Expected value (this picks out the center)). They summarize joint variation for random variables.
  - **Expected value** of a random variable X is the sum over all x in omega of little x * probability that X=little x. It is the weighted average of all possible outcomes. Do the f(x)*P(X=x) for each x and then add them.

    $$E(x) = \sum_{i=1}^{N} x_i p_i$$

    - To calculate expected value find x*P(X=x) for each x,
    - Then find f(x) for each x using the equation to define the relationship.
    - E(f(x)) = f(x)*P(X=x). This is the expected value of the new random variable.
- What happens when we take expected values of random variable and of functions of random variable
  - F(x) = function of future random variable.
    - X=price of milk
    - F(x) = units of milk to be sold.
    - If price of milk is random and quantity of milk is a function of price, then the quantity is also random.
    - Express expected value and variance in terms of properties of original random variables.
    - F could be a policy or decision.

- E(F(x)) = Sum over all possible outcomes of f(x)* probability that X=little x.

  - Expected value is special case where function is the identity function
  - Ex: X= stock price of Apple,
  - X P(X=x)
    100 .2
    150 .6
    200 .2
  - Expected value (F(x)) is sum of all possible outcomes: F(100)*.2 + F(150)*.6+F(200)*.2 =

- Rules for functions of random variables
  - Y = F (X), the function that is applied to random variable, is itself a new random variable that we will call Y. It is a linear function of random var. It is true that the expectation of Y is equal to A*Expectation of x. The expected value of function is the function of the expected value. Plug in expected value of x for x in the equation. The expected value of function of random variable, is equal to the funiton of the expected value.
    - E(Y) = E(F(X)) = a*E(x) + b
    - ie E(F(x)) = f(E(x)).
    - Can swap E and F, but only for linear functions. For all other functions, the expected value of a random variable is not equal to the function of the expected value.
    - Ex: X = price of milk, F(x) = 500*x^Beta; Beta = -2, F(x) is the demand curve. What is the expected value? (E(Q))
    - E(F(x)) = 64. So we expect to sell 64 units on average.
  - E(Q) = E(F(x)) not F(E(X))

x	P(X=x)	Expected Value	F(x)	F(x)*P(X=x)
2	0.2	0.4	125	25
3	0.6	1.8	55.55555556	33.33333333
4	0.2	0.8	31.25	6.25
		3		**64.58333333**
Given	Given	= x*P(x=x)	= 500*H2^-2	= F(x)*P(X=x)

Variance

The expected value of Z tell syou how far from the mean on average the random variable x is.

Y= F(X)

Z = [X – E(X)]^2

Mu = E(X)

Z = F(X) = (X-Mu)^2

E(Z) = variance (X)
Variance is measure of dispersion. Variance(X) = E {(X-E(X))^2}
= Sum for all of (X-Mu)^2* P(X-x)

Here Mu = Expected Value because he is using it as a short hand.

x	P(X=x)	Expected Value	X-E(x) or x-mu	(x-E(x))^2	(x-E(x))^2*P(X=x)	
2	0.2	0.4	-1	1	0.2	
3	0.6	1.8	0	0	0	
4	0.2	0.8	1	1	0.2	
		3			0.4	Sum of column
					0.632455532	Sqrt of .40

Joint Distribution / Multivariate Distributions

- Describes more interesting compete story about random variables that are correlated or have covariance.
- Table: if x and y can be only -1 or 1, this table shows all possible outcomes for joint probability. Brute force way to provide uncertainty about a random variable.

		Joint Probability
x	y	P(X=little X and Y=y)
-1	-1	0.3
1	-1	0.2
-1	1	0.2
1	1	0.3

All possible outcomes

- Variables:
  - X = % Return on Apple from now to next month
  - Y = % Return on Google
    - The above two are obviously correlated with each other.
- Some function f(X,Y): Could be a happiness function, policy decision, or etc.
- F(X.Y) is new random variable of policy of some future state of the world.
- E(F(X,Y)) = Sum of all states of the world (All possible joint events that incorporate both of what happens for X and Y. of F(xi,yi) * P(X=xi and Y=yi)
- Example: Simple Joint Distribution Data set.
  - F(X,Y) = .7X + .3(Y)
  - Apple is X and Google is Y
  - What is you expected Return: E(Z), where Z is the new random variable F(X,Y).
  - Step 1: Find(X,Y) for each X&Y combination.
  - Step 2: Multiply the Joint probability by each F(X,Y)

- Step 3: Sum this up.
  - Answer: 0.81
- Maximizing expected returns and minimizing variance is desired for a portfolio.
- 2 ideas we are developing with all of this:
  - What are the principles we should apply to these decision-making problems. (For stocks it is not clear do we want high expected value or low variance or somewhere in between)Expected Utility is a modification of expected value.
  - The other principle is what to do in real life. Need a better tool to address real life and all of its complexities. Bootstrap Resampling is a useful tool. We will go from 2 stocks to 500 stocks real quickly using this. This allows us to model much more probability distributions.