Optimization in Finance

Maximum Entropy Distribution of an Asset Inferrred from Option Prices

Bertrand Le Nezet

Numerical implementation steps

1. Option prices snap from Google Finance API
2. Data cleaning:
   - Calculate mid from bid / ask
   - Keep only relevant option prices (put-call parity/liquidity)
3. Apply the Cross Entropy Minimization algorithm
4. Compute option prices from asset distribution
5. Convert option prices into Black-Scholes Implied Volatility

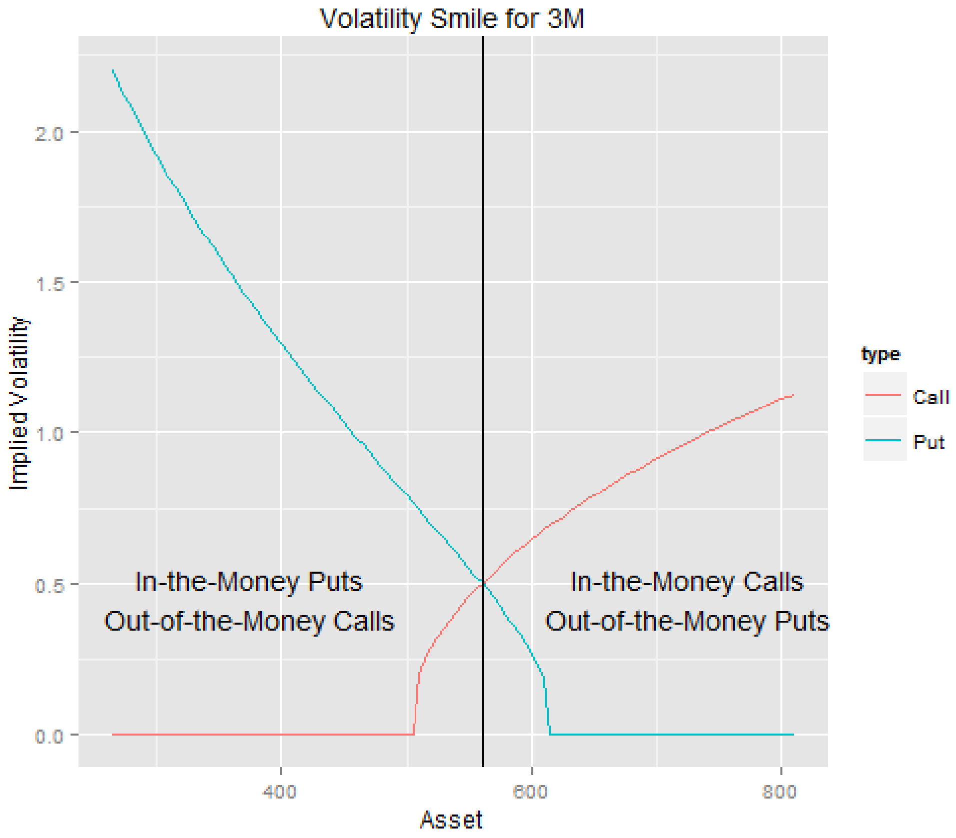
GOOG.OQ Option quotes

Google Inc (NASDAQ:GOOG)

View options by expiration Apr 24, 2015

Calls							Puts					
Price	Change	Bid	Ask	Volume	Open Int	Strike	Price	Change	Bid	Ask	Volume	Open Int
-	-	118.30	122.50	-	0	440.00	0.45	0.00	0.15	0.55	-	6
-	-	113.00	117.50	-	0	445.00	-	-	0.15	0.55	-	0
-	-	108.00	112.50	-	0	450.00	0.55	0.00	0.15	0.70	-	7
-	-	103.30	107.60	-	0	455.00	0.70	0.00	0.20	0.70	-	6
-	-	98.20	102.70	-	0	460.00	-	-	0.30	0.75	-	0
-	-	93.20	97.70	-	0	465.00	-	-	0.40	0.90	-	0
-	-	88.50	92.80	-	0	470.00	-	-	0.45	0.90	-	0
-	-	83.40	87.90	-	0	475.00	-	-	0.60	1.00	-	0
-	-	78.80	82.80	-	0	480.00	-	-	0.70	1.10	-	0
-	-	74.00	78.20	-	0	485.00	-	-	0.85	1.25	-	0
-	-	69.10	73.10	-	0	490.00	2.40	0.00	1.05	1.25	-	1
-	-	64.30	68.60	-	0	495.00	-	-	1.25	1.45	-	0
61.00	0.00	59.30	63.00	-	18	500.00	1.63	-0.57	1.50	1.75	1	2
-	-	51.60	54.20	-	0	510.00	2.35	-2.65	2.20	2.50	1	7
-	-	42.70	44.30	-	0	520.00	3.60	-0.90	3.20	3.70	7	9
-	-	40.10	42.10	-	0	522.50	-	-	3.60	4.00	-	0
-	-	38.50	40.00	-	0	525.00	7.22	0.00	4.00	4.40	-	3
-	-	35.90	37.90	-	0	527.50	-	-	4.40	4.90	-	0
33.70	0.00	34.40	35.90	-	9	530.00	7.40	0.00	4.80	5.40	-	107

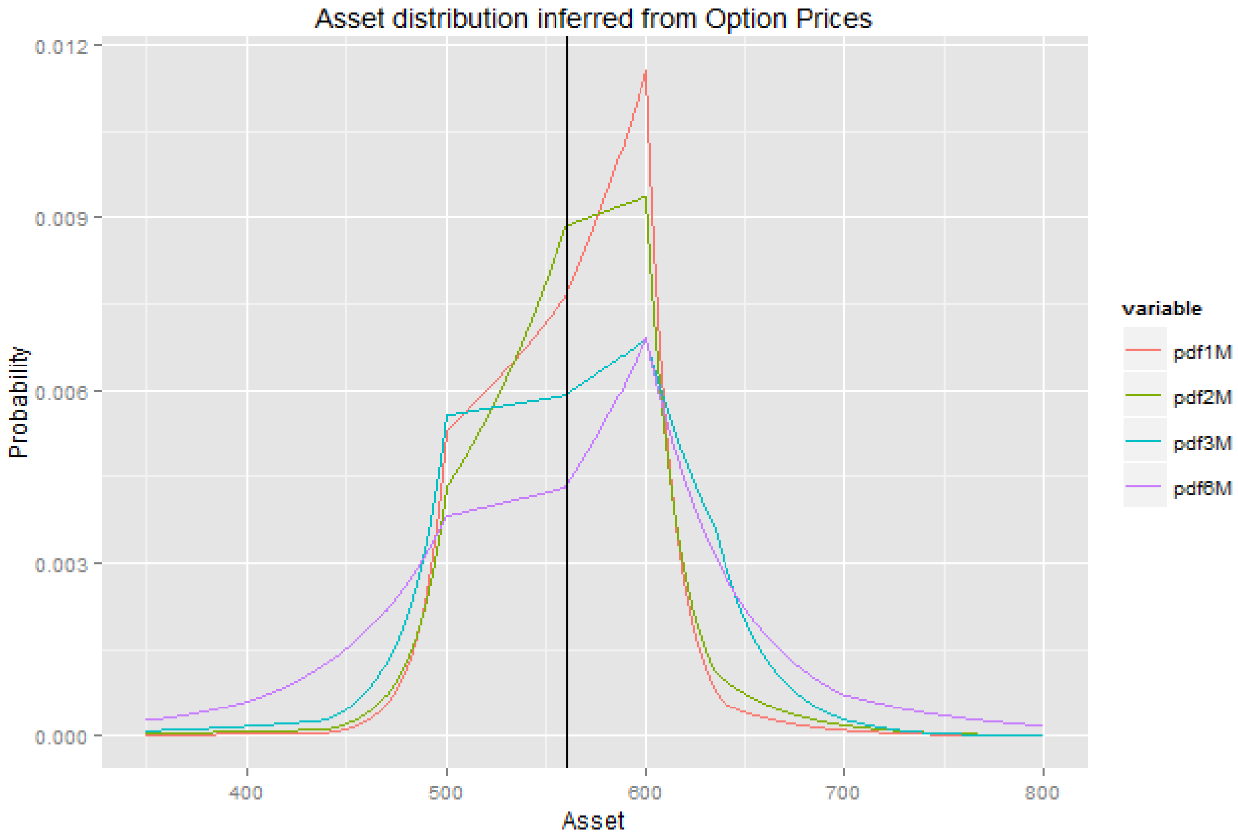
Volatility smile - Put-Call parity



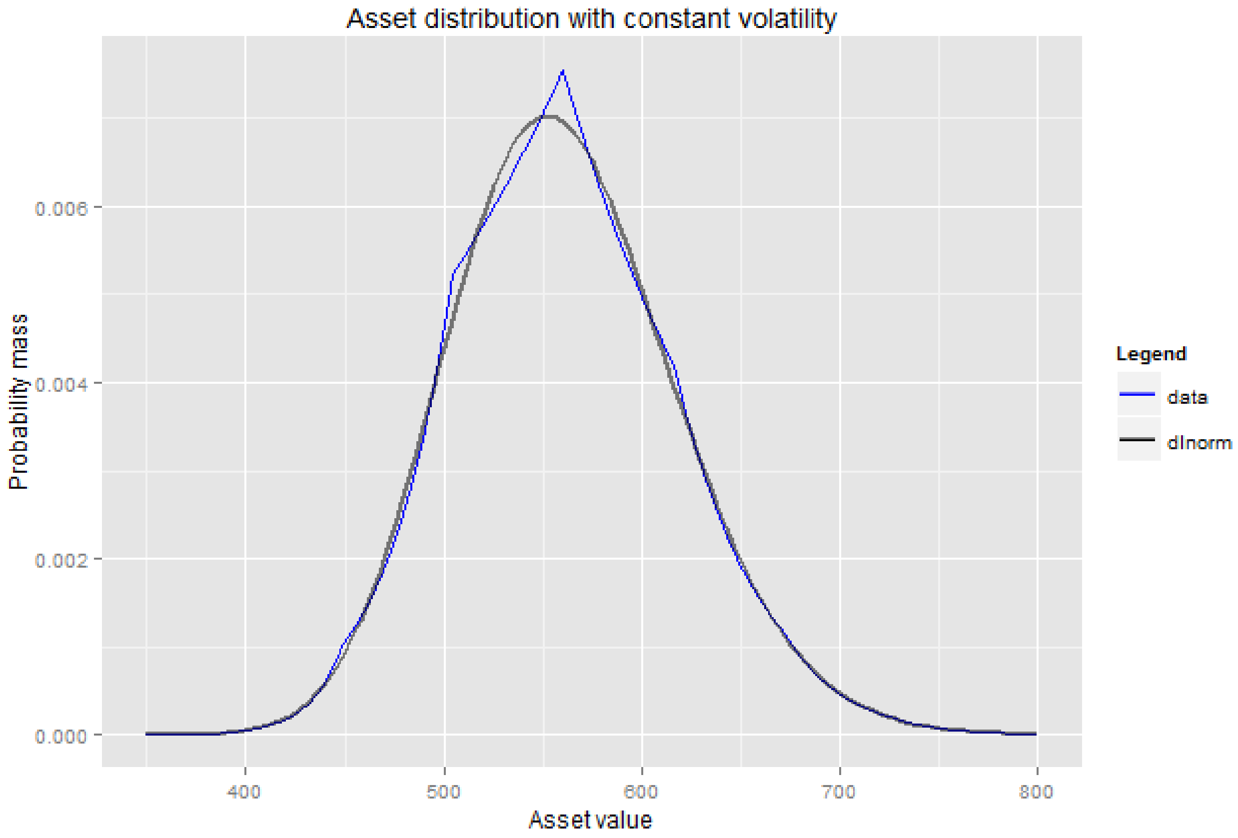
Cross Entropy Minimization algorithm

1. A uniform distribution is used as a prior (non-informed prior)
2. The option prices we snaped are used as constraint
3. The Lagrangian function (transforms constrained optimization into an unconstrainted optimization) is constructed
4. The Lagrangian function is minimized using an optimization routine (Limited-memory Broyden-Fletcher-Goldfarb-Shanno)

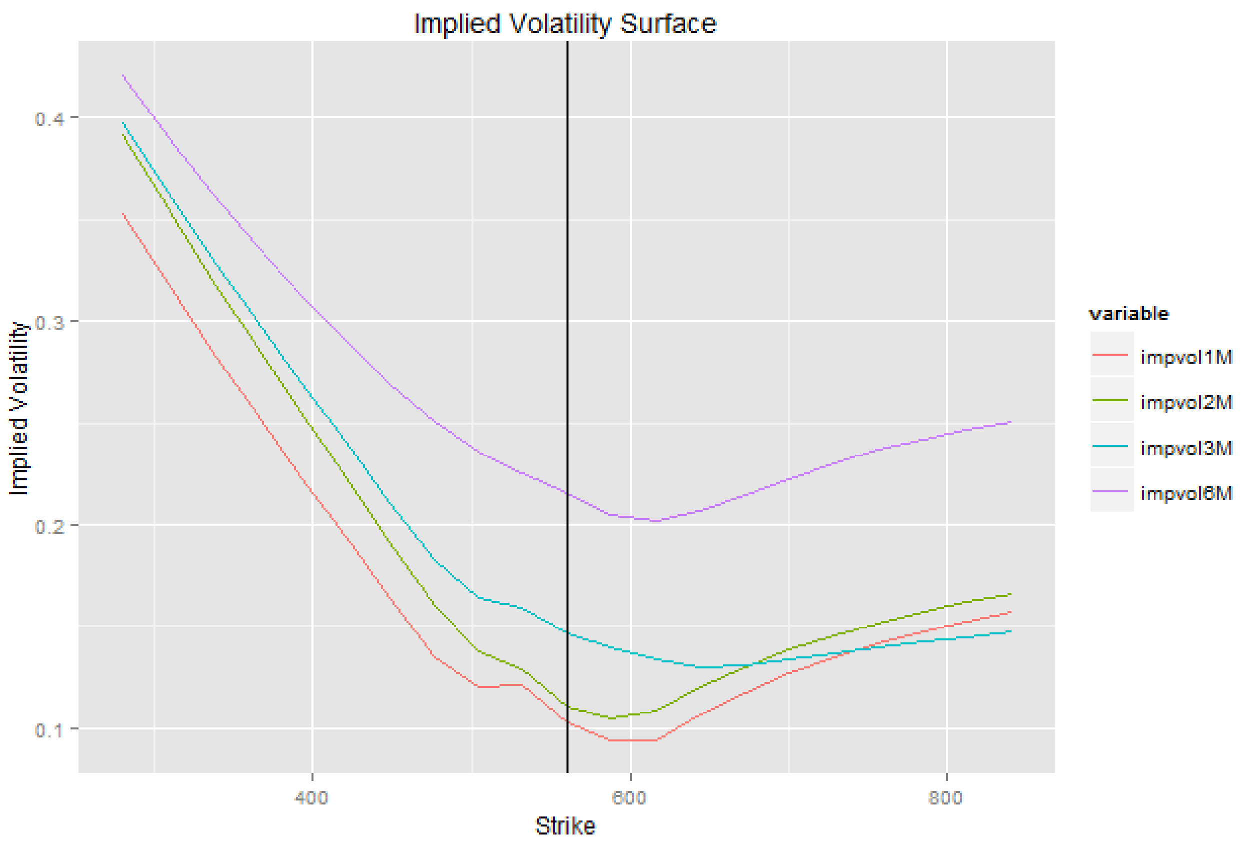
GOOG.OQ Asset distribution (1m/2m/3m/6m)



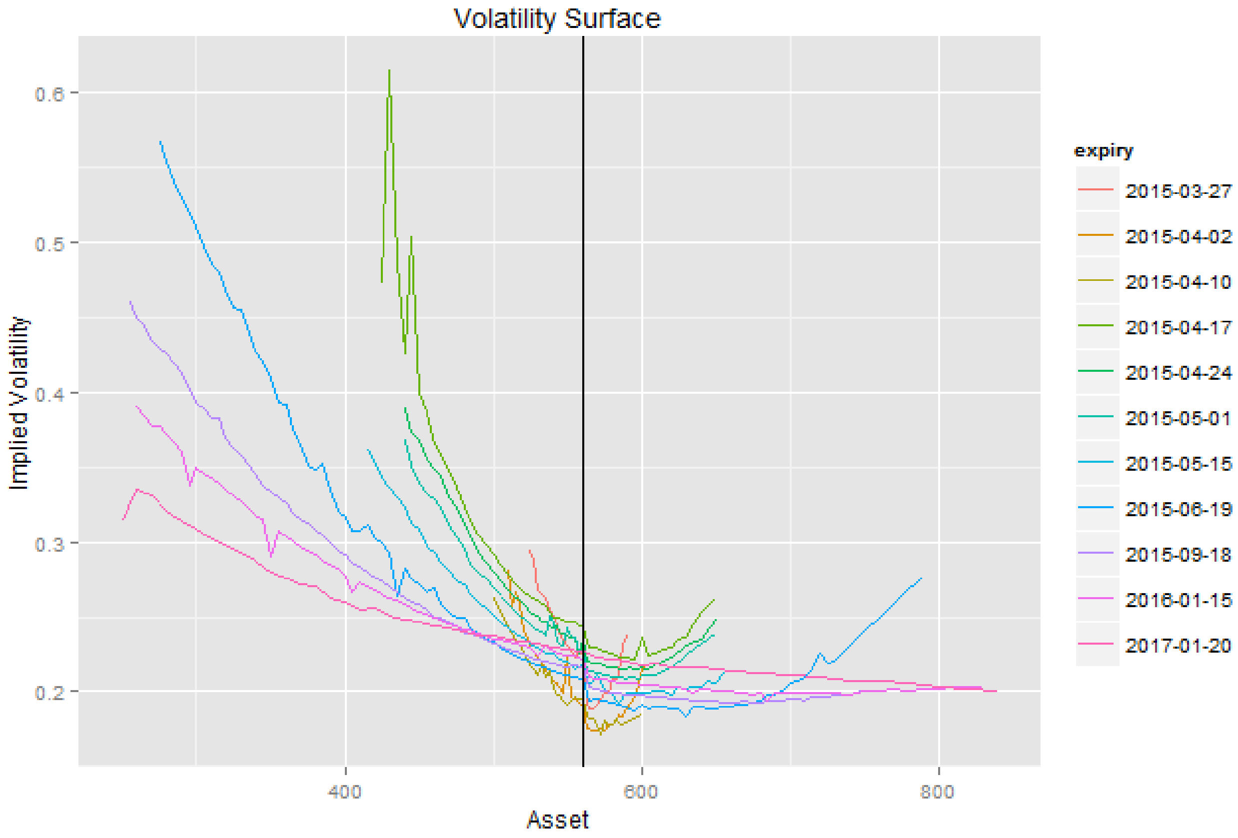
Asset distribution with constant volatility (25%)



GOOG.OQ Volatility smile (from Asset distribution)



Volatility smile (from Quotes)



Conclusion

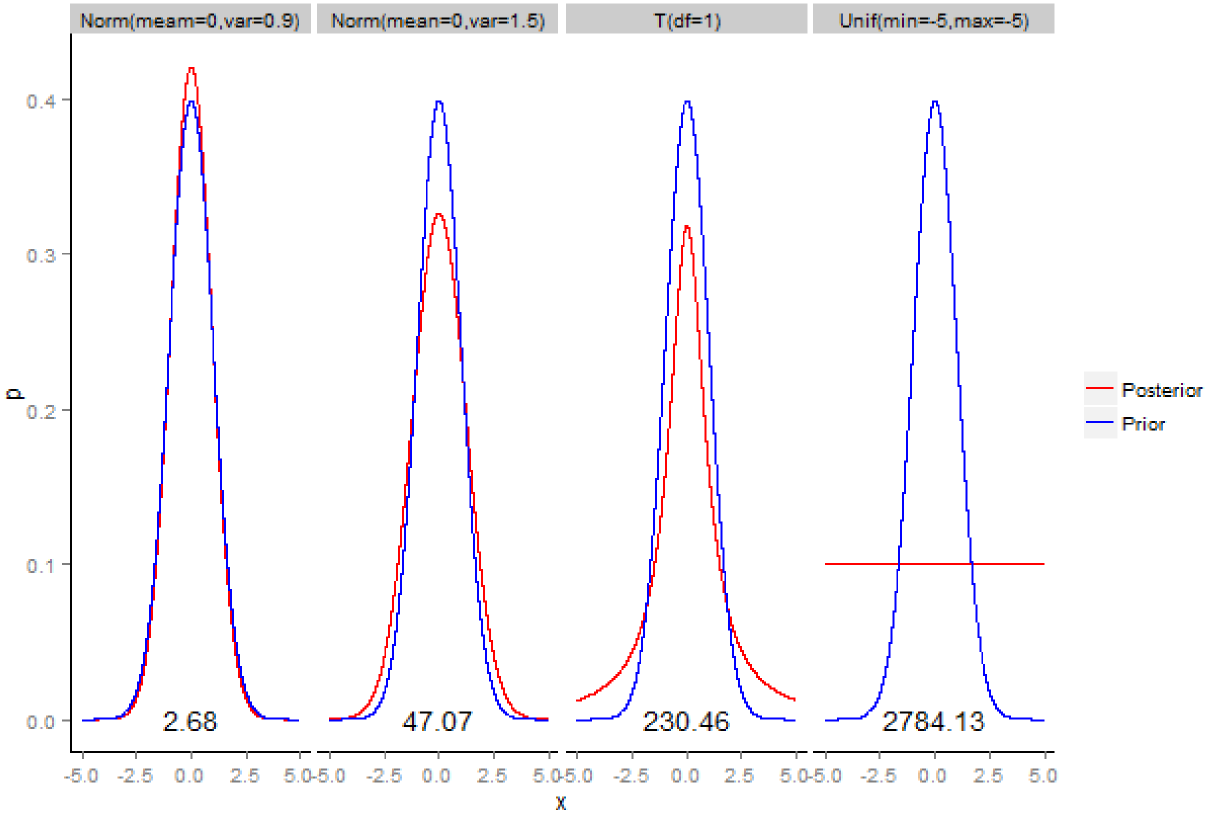
1. We have observed that assets distribution inferred from options prices are not log-normal (i.e. Volatility is not constant)
2. Principle of Minimum Cross-Entropy can be used to estimate the distribution of an asset without any assumption (non-parametric approach)
3. The algorithm is stable and fast
4. This method can be used to perform implied volatility interpolation/extrapolation from only few quotes

Appendix

Kullback-Leibler Relative entropy

1. In Information Theory, Shannon defines the entropy as a measure of unpredictability of information content
2. the KL relative entropy (or KL divergence) is a non-suymetric mesure of the difference between two probability distribution P and Q.
3. $D_{KL}(P\|Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, dx$
4. Properties:
5. KL is equal to zero if P and Q are identical
6. KL relative entropy is always positive

Calculation of Relative Entropy



Formulating the Optimization problem

We have one prior probability density function $q(x)$

We have m price constraints: $\forall i = 1, 2, ..m \; d_i = D(T)\mathbb{E}_{\mathbb{Q}}[c_i(X_T)]$where

- $D(T) = e^{-(r-q)T}$ represents the discount factor
- $c_i(X_t)$ denotes the ith option pay-off function at expiry depenent only on the asset value at expiry
- d_i is the corresponding option price
- r risk-free rate for T
- q dividend yield

Formulating the Optimization problem (2)

Minimize $S(p,q) = \int_{-\infty}^{\infty} p(x) \log\left[\frac{p(x)}{q(x)}\right] dx$

Subject to 2 constraints:

1. $\int_0^{\infty} p(x)dx = 1$
2. $\forall i = 1, 2, .. m \; \mathbb{E}[c_i(X)] = \int_0^{\infty} p(x) \; c_i(x)dx = ci$

This is a standard constrained optimization problem which solved by using the method of Lagrange which transforms a problem in n variable and m constraints into an unconstrainted optimization with n+m variables.

Objective function

$$H(p) = -\int_0^\infty p(x)\log\left[\frac{p(x)}{q(x)}\right]dx + (1+\lambda_0)\int_0^\infty p(x)dx + \sum_{i=1}^m \lambda_i \int_0^\infty p(x)\; c_i(x)dx$$

From standard calculus, we know that the minimum $\lambda^* = (\lambda_0^*, \ldots, \lambda_M^*)$ is reached when:

- the gradient (vector of derivatives) δH is equal to zero: $\delta H(\lambda^*) = \int_0^\infty \left[-\log\left[\frac{p(x)}{q(x)}\right] + \lambda_0 + \sum_{i=1}^m \lambda_i c_i(x)\right]\delta p(x)dx = 0$ (necessary condition)
- the hessian (matrix of second derivatives) is positive definite (sufficient condition)

Objective function solution

This leads immediately to the following explicit representation of the MED:

$p(x) = \frac{q(x)}{\mu} \exp\left(\sum_{i=1}^{m} \lambda_i c_i(x)\right)$, $\mu = \int_0^\infty q(x) \exp\left(\sum_{i=1}^{m} \lambda_i c_i(x)\right) dx$

Impact of constraints on MXED