Random Variables in EU

- Specification of the random variables can universally be summarized by their joint probability distribution.

- The form of the distributional information has a large impact on model formulation:
  - Continuous
  - Discrete
  - Mixed
  - Moments

Random Variables (cont’d.)

- Continuous random variables are specified through a density function and a domain of integration.
  - E.g., the joint normal density is

\[ f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} \left| C \right|^{-1/2} \exp \left( -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right) \]

where $C$ is a constant, and its domain is all of $n$-dimensional space.
Random Variables (cont’d.)

- Discrete distributions are specified via a list of discrete points and probabilities \((\beta^k, p^k), k = 1, \ldots, K\)

  - E.g., the binomial distribution is

    \[
    p^k = \binom{K}{k} \delta^k (1 - \delta)^{K-k}, \quad k = 0,1,\ldots, K
    \]

    where \(\delta\) is a fixed value strictly between zero and one (Note that these probabilities do sum to 1.0!)

- If multivariate, these tend to be empirical distributions

Random Variables (cont’d.)

- Mixed distributions are specified as a mixture of a continuous and a discrete distribution

  - E.g., We could define a “truncated normal” distribution as having the univariate standard normal distribution for \(\beta > 0\), and having “point mass” at \(\beta = 0\) (What does the CDF look like?)

  - These are quite uncommon as the fundamental variables

  - Some payoff functions are similar (e.g., insurance)
Truncated Normal with Point Mass CDF

- CDF of a normal truncated from below with point mass at truncation point

Random Variables (cont’d.)

- A limited number of moments may be all that is available in some cases
  - E.g., perhaps only the mean vector and covariance are available
  - Generally, more moments are better than fewer, but one must pay attention to the ultimate use in the math program (worry about convexity/concavity!)
  - Gaussian quadrature and entropy may play a role
Random Variables (cont’d.)

- Gaussian quadrature and entropy – an aside

- At times, one needs to have a discrete approximation to a distribution when only moments are available

- Gaussian quadrature produces such an approximation that contains the fewest points (and least uncertainty)

- Entropy produces such an approximation that contains the most points (and most uncertainty)

Random Variables (cont’d.)

- For either problem,

  - The first step is to generate a bunch of points in the domain for the random variables $x_k$

  - The second step is to compute probabilities that satisfy some “moment constraints” that extremize some objective

- The difference between the two methods has to do with the choice of objective
Random Variables (cont’d.)

- With two random variables, moment constraints are:
  \[
  \sum_{k=1}^{K} (\beta_1^k)^j (\beta_2^k)^j p^k = E[\beta_1^j \beta_2^i], \quad i, j = 0, ..., M; \ i + j \leq M
  \]

- In addition, we require that \( p^k > 0 \), and that these probabilities sum to 1.0

- For Gaussian quadrature, the objective is vacuous (there is none)
  \[
  \sum_{k} p^k \leq + \sum_{k} (\beta_1^k)^j (\beta_2^k)^j p^k \]

Random Variables (cont’d.)

- For maximum entropy, the objective is:
  \[
  \text{maximize} \quad - \sum_{k} p^k \ln(p^k)
  \]

- How many probabilities will be positive now?
  - Hint: look at the marginal entropy as some \( p^k \) approaches zero.
  \[
  \lim_{p^k \to 0} - \frac{\partial}{\partial p^k} \ln(p^k) = \lim_{-1 - \ln(p^k)} = -\infty
  \]
Payoff Function

- The payoff function is entirely case specific and its construction employs a combination of
  - Accounting (to get the data)
  - Economics (to get appropriate behavioral implications of the data)

Implementations of Risky Choice Models

- A large number of alternative implementations of models of choice under risk appear in the literature
- The primary drivers of the choice of approach have historically been:
  - The form of the distributional information on the random variables
  - Expertise and software availability for the class of problem implied by the model
Risk Model Implementations

The primary difficulty with this type of model is the objective function

Continuous case:

\[ E[u(W(x, \beta))] = \int u(W(x, \beta)) f(\beta) d\beta \]

- Notice that this is a multi-dimensional integral

\[ f(\beta) \] denotes the joint distribution of the random variables

Implementation (cont’d.)

- The integral will often have no closed form solution

- Even if the integral does have a closed form solution, the resulting objective will be nonlinear

Most implementations of the above problem are based on:

- Simplifications resulting from distributional choices

- Approximations to the expected utility function

- Approximations of the joint distribution
Implementation (cont’d.)

Discrete case:

\[ E[u(W(x, \beta))] = \sum_{i=1}^{n} u(W(x, \beta_i)) p_i \]

- In this case, the subscripts \( i \) do not indicate elements of the beta vector, but discrete realizations of the random variables (e.g., head or tail)

- \( \beta_i \) denotes the vector of outcomes in “state of nature” \( i \)

- \( p_i \) denotes the probability of state of nature \( i \)

Implementation (cont’d.)

- This formulation remains nonlinear, but no longer requires a closed form expression for a multi-fold integral

- This is the formulation that results if we make a discrete approximation to the joint distribution

- This is a handy formulation when you want to reflect data directly in the problem -- if the \( \beta \) are observations from the real world, and \( p_i = 1/n \), then we call this an empirical distribution
Implementation (cont’d.)

For illustration, a variety of implementations of a particularly simple but theoretically important model of choice under risk -- the single-period portfolio model

\[
\begin{align*}
\text{maximize} & \quad E[u(W(x, \beta))] = E[u(\beta'x)] \\
\text{subject to} & \quad \sum_{i=1}^{n} x_i \leq I
\end{align*}
\]

- \(x_i\) denotes the amount of initial wealth, \(I\), invested in asset \(i\)
- The payoff function is bilinear in the random variables and the choice variables

The Mean/Variance or E/V model

\[
\begin{align*}
\text{maximize} & \quad E[\beta']x - \frac{\rho}{2} x' E[\beta - E[\beta]]E[\beta - E[\beta]']x \\
\text{subject to} & \quad \nabla x_i \leq I
\end{align*}
\]

- Freund (1956) showed that if the beta’s are distributed as joint normal random variables and if the utility function is a risk averse (negative) exponential then this problem produces identical results to the portfolio problem where rho is the coefficient of absolute risk aversion
Implementation (cont’d.)

- Meyer showed that if the beta’s are joint normal, then there exists a rho that yields the optimum of the portfolio problem — regardless of which risk averse utility function is the true one

- Notice the differences in these statements
  - Freund shows equivalence between E/V and the portfolio problem
  - Meyer shows that every optimum to a portfolio problem is optimal for the E/V problem with some value for rho

Note that the above problem is a quadratic programming problem

In Freund’s time, these problems were hard

Subsequent to Freund, a number of linear programming approximations to the expected utility (or E/V) problems were developed
Implementation (cont’d.)

I MOTAD (Hazell, 1971) or Mean Of Total Absolute Deviation -- One of the best known of the linear programming approximations to the E/V model

I Assumptions

• The payoff function is bi-linear in the random variables and choice variables (as in the portfolio problem)

• The states of nature are equally likely (e.g., as with an empirical distribution)

Implementation (cont’d.)

I Motivation

• Let \( \bar{\beta} = \frac{1}{n} \sum_{i=1}^{J} \beta^i \) denote the mean of random variable \( i \)

• \( E[W(x, \beta)] = \bar{\beta}^\prime x = \sum_{i=1}^{n} \bar{\beta}^i \bar{x}^i \) and

• \( ar(W(x, \beta)) = \frac{1}{n} \sum_{i=1}^{J} \left( \sum_{i=1}^{n} x^i (\beta^i - \bar{\beta})^2 \right) \)
Implementation (cont’d.)

- The idea with MOTAD is that we approximate the above variance with the following:

\[ ar(W(x, \beta)) = \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{n} x_i (\beta_{ij} - \bar{\beta}_i) \]

- What part of the variance is being approximated with the absolute value?

- Why is this similar to a variance?

- How is this different from the variance?

Implementation (cont’d.)

- Notice that this measure of variance cannot be incorporated directly into the objective because it is not differentiable.

- These non-differentiabilities can be avoided by introducing additional constraints (equalities and inequalities) and introducing additional variables.
Implementation (cont’d.)

- To incorporate the absolute value based approximation to variance, we first introduce some variables and constraints
  
  \[ < \text{some deviation} > -d^+ + d^- = 0 \]

- We restrict the new variables to be non-negative

Implementation (cont’d.)

- Consider the optimum to this problem:

  \[
  \min_{d^+, d^- \geq 0} d^+ + d^- \\
  \text{subject to:} \quad + -
  \]
Implementation (cont’d.)

1. The MOTAD formulation:

\[
\text{minimize } \sum_{j=1}^{n} (d_{j}^{+} + d_{j}^{-})
\]

subject to:

\[
\sum_{j=1}^{n} x_{j} \beta_{j} \geq \mu \\
\sum_{j=1}^{n} x_{j} (\beta_{ij} - \beta_{j}) - d_{j}^{+} + d_{j}^{-} = 0 \\
\sum_{j=1}^{n} x_{j} \leq I, \ x_{j} \geq 0, \ d_{j}^{+} \geq 0, \ d_{j}^{-} \geq 0.
\]

Implementation (cont’d.)

- Notice that we are now minimizing the approximation to variance subject to “doing well enough” on the mean.

- Does this problem produce the same optimal \( x_{i} \) values as if we were minimizing the MOTAD measure of variance?

- Where is the risk aversion level?

- How does this relate to the E/V problem?

1. Problem -- this approach penalizes deviations above and below the mean equally.
Implementation (cont’d.)

- Target MOTAD (Tauer, 1983)
  - Similar to MOTAD, but does not penalize deviations above “target”
  - Requires user to specify a target as well as a mean level of income

Formulation of Target MOTAD:

\[
\text{minimize } \sum_{n} d_{ij} \geq 0
\]

subject to:

\[
\sum_{n} x_{i} \beta_{j} \geq \mu
\]

\[
\sum_{n} x_{i} \beta_{j} - T + d_{ij} \geq 0
\]

\[
x_{i} \leq I, \ x_{i} \geq 0, \ d_{ij} \geq 0.
\]
Implementation (cont’d.)

### Other linear programming based models of choice under risk

- MaxiMin Criterion
- Safety-first models
- Etc.

These tend to focus on extremes -- i.e., what happens in the worst case

-------

Implementation (cont’d.)

### Mean/Variance Models -- Two Formulations

\[
\begin{align*}
\text{maximize} & \quad \sum_{j} W_j - \frac{\rho}{2} \sum_{j} \left( W_j - \sum_{k} W_k \right)^2 \\
\text{subject to} & \quad \sum_{j} x_j \beta_j = W_j \\
& \quad v \leq I, \quad v \geq 0.
\end{align*}
\]

-------
Implementation (cont’d.)

1. Formulation 2

maximize \( \frac{1}{J} \sum_{j=1}^{J} \sum_{i=1}^{n} x_i \beta_{ij} \)

\[ - \frac{\rho}{2J} \sum_{j=1}^{J} \left( \sum_{h=1}^{n} \sum_{i=1}^{n} x_i (\beta_{ij} - \bar{\beta}_i) (\beta_{hj} - \bar{\beta}_h) x_h \right) \]

subject to: \( \sum_{i=1}^{n} x_i \leq I, \quad x_i \geq 0, \)

or in matrix notation

maximize \( \beta' x - \frac{\rho}{2} x' E[(\beta - \bar{\beta})(\beta - \bar{\beta})'] x \)
Implementation (cont’d.)

- Notice that we don’t need to know the joint distribution of the betas, only the mean and covariance matrix.

- Thus, this approach can be applied in the continuous or discrete cases, or even when only the moments (not the distribution) are known.

- All of the E/V problems above are (strictly convex) quadratic programming problems for risk averse agents.

Implementation (cont’d.)

- Direct Expected Utility Models (Lambert and McCarl, 1985)

- Based on directly implementing the expected utility model.

- Joint distribution of random variables must either be discrete or approximate as discrete.
Implementation (cont’d.)

1 Formulation of portfolio model:

\[
\text{maximize } \sum_{j=1}^{J} p_j u(W_j)
\]

subject to:

\[
\sum_{i=1}^{n} x_i \beta_{ij} = W_j
\]

\[
\sum_{i=1}^{n} x_i \leq 1, \ x_i \geq 0.
\]

Implementation (cont’d.)

1 Or more generally:

\[
\text{maximize } \sum_{j=1}^{J} p_j u(W(x, \beta_j))
\]

subject to:

\[
\sum_{i=1}^{n} x_i \leq 1, \ x_i \geq 0.
\]

1 Generalization of each of these problems to the case of more general constraints (than the portfolio problem) is straightforward.