Probability in

OPTIMIZATION MODELS

Joydeep Dutta
Dept of Economic Sciences,
IIT Kanpur.

Portfolio and returns

Portfolio : A collection of assets in which investment is made

We will consider a portfolio with n-assets.

Total Return R:

$$R = \frac{X_1}{X_0} = \frac{\text{Amout recieved}}{\text{Amount Invested.}}$$

Practically important: Relative Return

$$r = \frac{X_1 - X_0}{X_0}$$

So $$\boxed{R = 1 + r}$$ → Acts like short term rate of interest.

Further $$\boxed{X_1 = X_0(1+r)}$$

- In our market model short-selling is allowed.

Investment in asset i initially

$$\boxed{X_{0i} = \omega_i X_0}$$

→ Weights (Decision variable)

Relative return on asset i

$$\boxed{r_i = \frac{X_{1i} - X_{0i}}{X_{0i}}}$$

A random variable for each asset i

So for the portfolio, has a return random vector

$$\vec{r} = (r_1 \ldots\ldots r_n)^T$$

Relative return r of the portfolio

$$r = \frac{X_1 - X_0}{X_0}$$

$$= \frac{1}{X_0}\left(\sum_{i=1}^{n} X_{1i} - \sum_{i=1}^{n} X_{0i}\right)$$

$$= \frac{1}{X_0}\left(\sum_{i=1}^{n} (1+r_i)X_{0i} - \sum_{i=1}^{n} X_{0i}\right)$$

$$\boxed{r = \sum_{i=1}^{n} \omega_i r_i = \langle \omega, \vec{r} \rangle}$$

random variable too.

$$E(r) = \sum_{i=1}^{n} \omega_i E(r_i)$$

You can safely assume that they follow $N(\mu_i, \sigma_i^2)$

$$Var(r) = E\left(r - E(r)\right)^2$$

$$= E\left(\sum_{i=1}^{n} \omega_i r_i - \sum_{i=1}^{n} \omega_i E(r_i)\right)^2$$

If $E(r_i) = \bar{r}_i$, then

$$Var(r) = E\left(\sum_{i=1}^{n} \omega_i (r_i - \bar{r}_i)\right)^2$$

$$= \sum_{i,j} \omega_i \omega_j \, E(r_i - \bar{r}_i)(r_j - \bar{r}_j)$$

$$= \sum_{i,j} \omega_i \omega_j \sigma_{ij} \quad [\sigma_{ij} = Cov(r_i, r_j)].$$

$$\boxed{\sigma_{ii} = \sigma_i^2 = Var(r_i).}$$

Markowitz viewed $Var(r)$ as a <u>measure of risk</u>.

<u>The Markowitz Portfolio model</u> (Mean-variance model)

$$\boxed{\begin{aligned} &\text{minimize } Var(r) \\ &\qquad E(r) = \rho \\ &\qquad \sum_{i=1}^{n} \omega_i = 1 \end{aligned}}$$

Markowitz got the Nobel memorial prize in Econ. (1990).

<u>Level</u> of <u>return desired</u>

This is a <u>parametric stochastic optimization problem</u>.

$$\mathrm{Var}(r) = \langle \omega, \Sigma \omega \rangle \quad \text{where}$$

$$\Sigma = [\sigma_{ij}]_{n \times n}$$

is the variance - covariance matrix. This is a symmetric and positive semi-definite matrix.

- Key assumptions for solving the model

A) Σ is positive definite

B) All r_i's dont have the same value.

An optimization detour

The Markowitz problem has a solution and that satisfies a KKT condition / Lagrange multiplier rule. That is as follows. There exists scalars λ and μ, such that

$$\sum_{j=}^{n} \sigma_{ij} \omega_j - \lambda \bar{r}_i - \mu = 0, \quad i = 1, 2, \ldots n$$

$$\sum_{i=1}^{n} \omega_i \bar{r}_i = \rho$$

$$\sum_{i=1}^{n} \omega_i = 1$$

(For the Econ Folks)

Let us write $\bar{r} = (\bar{r}_1, \bar{r}_2 \dots \bar{r}_n)$; then the problem is stated as

$$\boxed{\begin{array}{l} \min \frac{1}{2} \langle \omega, \Sigma \omega \rangle \\ \text{Subject to} \\ \qquad \langle \omega, \bar{r} \rangle = \rho \\ \qquad \langle e, \omega \rangle = 1 \end{array}}$$

where $e = (1, 1, \dots 1)^T$

Based on assumptions A) and B) we shall solve it.

By A), Σ is positive definite. Hence Σ^{-1} exists.

Set

$$\boxed{\begin{array}{ll} \alpha = \langle e, \Sigma^{-1} e \rangle & \beta = \langle e, \Sigma^{-1} \bar{r} \rangle \\ \gamma = \langle \bar{r}, \Sigma^{-1} \bar{r} \rangle & \delta = \alpha\gamma - \beta^2 \end{array}}$$

The Lagrangian

$$L(\omega, \lambda, \mu) = \frac{1}{2} \langle \omega, \Sigma \omega \rangle - \lambda \left(\langle \bar{r}, \omega \rangle - \rho \right) - \mu \left(\langle e, \omega \rangle - 1 \right).$$

The KKT conditions again

$$\Sigma \omega - \lambda \bar{r} - \mu e = 0$$

Thus

$$\boxed{\omega = \Sigma^{-1}(\lambda \bar{r} + \mu e)}$$

We need to find λ and μ.

$$\langle \bar{r}, \omega \rangle = \rho$$

$$\langle \bar{r}, \Sigma^{-1}(\lambda \bar{r} + \mu e) \rangle = \rho$$

$$\lambda \langle \bar{r}, \Sigma^{-1} \bar{r} \rangle + \mu \langle \bar{r}, \Sigma^{-1} e \rangle = \rho$$

$$\boxed{\lambda \gamma + \mu \beta = \rho}$$

Now

$$\langle e, \omega \rangle = 1$$

$$\langle e \quad \Sigma^{-1}(\lambda \bar{r} + \mu e) \rangle = 1$$

$$\lambda \langle e, \Sigma^{-1} \bar{r} \rangle + \mu \langle e, \Sigma^{-1} e \rangle = 1$$

$$\boxed{\lambda \beta + \mu \alpha = 1}$$

Thus
$$\boxed{\mu = \frac{\gamma - \beta\rho}{\delta} \quad \& \quad \lambda = \frac{\alpha\rho - \beta}{\delta}}$$

That solves it completely and gives

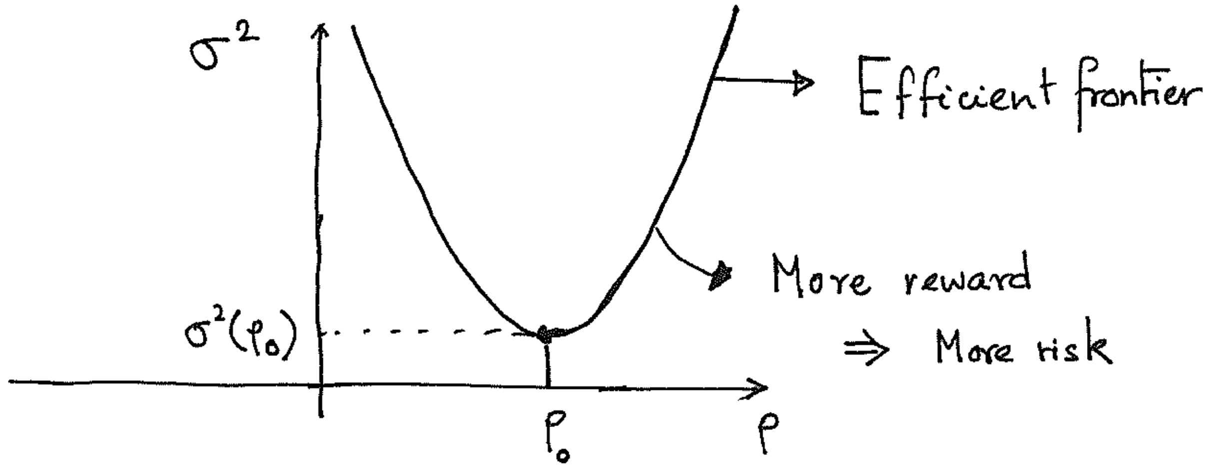
$$\sigma^2(\rho) = \langle \omega, \Sigma\omega \rangle$$

by convexity, as ω satisfies KKT, it is the optimal.

$$\Rightarrow \boxed{\sigma^2(\rho) = \frac{\alpha\rho^2 - 2\rho\beta + \gamma}{\delta}}$$

Strongly convex

$\alpha > 0$, $\gamma > 0$. Thus we have the following



General Structure of a Stochastic Optimization

$$\boxed{f(x, \xi)}$$ → Objective function model

x → decision

ξ → Uncertain Parameter

Uncertain Parameter → Robust Optimization

Uncertain Parameter → Stochastic Optimization

The Robust formulation

Let U be the uncertainity set. Then consider the worse case scenario

$$\max_{\xi \in U} f(x,u)$$

Robust problem: $$\boxed{\min_{x \in X} \max_{\xi \in U} f(x,u)}$$

Linked to bilevel programming

The stochastic formulation

$(\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow$ given probability space

and $\xi : \Omega \longrightarrow \mathbb{R}^m$ be a random vector.

$\mathbb{R}^m$ equipped with Borel σ- algebra

$$\boxed{\min_{x \in X} E_{\mathbb{P}} [f(x, \xi)]}$$

Because ξ depends on ω, we can write

$$f(x, \xi) = F(x, \omega)$$

Thus $$\boxed{\min_{x \in X} E_{\mathbb{P}} [\cancel{f(x,} F(x, \omega)]}$$

Note that

$$\boxed{E_{\mathbb{P}}\left[f(x,\xi)\right] = \int_{\Omega} f(x,\xi)\, d\mathbb{P}}$$

We will define the expectation functional

$$\phi : x \longmapsto E_{\mathbb{P}}[f(x,\xi)]$$

$$\phi(x) = E_{\mathbb{P}}[f(x,\xi)]$$

Let us try to learn a bit more about the expectation functional.

A) Finite Distributions

A random vector ξ is finitely distributed if ξ takes finite number of values. Let $\xi : \Omega \longrightarrow \mathbb{R}^m$. Let $\hat{U} \subseteq \mathbb{R}^m$ be the set containing the values of ξ. Support of ξ, call $\boxed{\operatorname{supp}\xi = \{ p_\xi > 0,\ \xi \in \hat{U} \}}$

The distribution function

$$F(\eta) = \sum_{\xi \leq \eta} p_\xi \quad \forall\, \eta \in \mathbb{R}^n \left[\sum_{\xi \in \operatorname{supp}\hat{U}} p_\xi = 1 \right]$$

$$\boxed{\phi(x) = E_{P}\left[f(x,\xi)\right] = \sum_{\xi \in \hat{U}} f(x,\xi)\, p_\xi}$$

B) <u>Discrete</u> <u>Distribution</u>

The random vector has ~~a~~ countable number of values, i.e.

$$\hat{U} = \{ \xi_1, \xi_2, \xi_3 \ldots \xi_n \ldots\ldots \}$$

$\therefore$ Here

$$E_{\mathbb{P}}\left[f(x,\xi)\right] = \sum_{\xi \in \hat{U}} p_\xi \, f(x,\xi)$$

of course $\sum_{\xi \in U} p_\xi = 1.$

Key issue here $\boxed{\sum_{\xi \in \hat{U}} p_\xi \, f(x,\xi)}$ → need not be finite (ie need not converge)

<u>Key</u> <u>assumption</u> $\boxed{\infty - \infty = \infty}$

$$\boxed{E_{\mathbb{P}}\left[f(x,\xi)\right] = E_{\mathbb{P}}\left[\max\{f(x,\xi), 0\}\right] - E_{\mathbb{P}}\left[\max\{-f(x,\xi), 0\}\right]}$$

c) <u>Continuous Case</u>

$$E_{\mathbb{P}}\left[f(x,\xi)\right] = \int_{\mathbb{R}^m} \max\{f(x,\xi), 0\}\, p(\xi)\, d\xi - \int_{\mathbb{R}^m} \max\{-f(x,\xi), 0\}\, p(\xi)\, d\xi$$

<u>probability ~~distribution~~</u> <u>density</u> of :

Convexity, Continuity and Differentiability

- $f(x, \xi)$ is integrable for all $x \in \mathbb{R}^n$
- $f(x, \xi)$ is convex for all $\xi \in \mathbb{R}^m$

$\Big\}$ given

Then $\phi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is well-defined and convex.

Here $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

$$\phi(x) = E_{\mathbb{P}}[f(x, \xi)]$$

Thus $\phi(\lambda y + (1-\lambda)x)$

$$= E_{\mathbb{P}}[f(\lambda y + (1-\lambda)x, \xi)] \qquad (\lambda \in [0,1])$$

By convexity

$$f(\lambda y + (1-\lambda)x, \xi) \leq \lambda f(y, \xi) + (1-\lambda) f(x, \xi)$$

$$E_{\mathbb{P}}[f(\lambda y + (1-\lambda)x, \xi)] \leq E[\lambda f(y, \xi) + (1-\lambda) f(x, \xi)]$$

$$= \lambda E[f(y, \xi)] + (1-\lambda) E[f(x, \xi)]$$

Voila : We are done!!

\# Continuity

\# Differentiability

Continuity of the Expectation Functional

Let $f(\cdot, \xi)$ is continuous at x_0 for each $\xi \in \mathbb{R}^m$.
Assume that $f(\cdot, \xi(\omega))$ is measurable for every x in a neighborhood of x_0. There exists a measurable function $G : \Omega \to \mathbb{R}$ such that $E_{\mathbb{P}}[G] < +\infty$ and

$$|f(x, \xi(\omega))| \leq G(\omega), \quad \mathbb{P}\text{- a.e.}$$

for any x in a neighborhood of x_0, then the expectation functional is also continuous at $x = x_0$

Proof: Consider $x^k \to x_0$, by continuity,

$$\lim_{k \to \infty} f(x_k, \xi(\omega)) = f(x_0, \xi(\omega))$$

Now by the Dominated Convergence Theorem

$$\lim_{k \to \infty} \int_{\Omega} f(x_k\ \xi(\omega))\, d\mathbb{P} = \int_{\Omega} f(x_0, \xi(\omega))\, d\mathbb{P}$$

$$\boxed{\lim_{k \to \infty} \phi(x_k) = \phi(x_0)}$$

Note

$$\phi(x_k) = E_{\mathbb{P}}[f(x_k, \xi)] = \int_{\Omega} f(x_k, \xi(\omega))\, d\mathbb{P}.$$

Hence the result.

Computing differentials / Rather Subdifferentials

Will focus on the case when $f(\cdot, \xi)$ is convex.

Can we compute $\partial\phi(x)$. The subdifferential of the expectation functional

Basics: Set-Valued Maps

$S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is a set-valued map

$$\boxed{\xi \in \mathbb{R}^m \longmapsto S(\xi) \subset \mathbb{R}^n}$$

$$\text{dom } S = \{\xi \in \mathbb{R}^m : S(\xi) \neq \emptyset\}$$

If $\forall\ \xi$, $S(\xi)$ is a convex set, then S is convex-valued and if $S(\xi)$ is a compact set, then S is compact-valued

$$\boxed{S(\xi) = \partial_x f(\bar{x}, \xi)}$$

Subdifferential

Convex valued   Compact valued

Random Sets: The Key notion to define a random set is the notion of "measurable multifunction"

Let $\xi : \Omega \rightarrow \mathbb{R}^m$, and let $\xi(\Omega) = U \subseteq \mathbb{R}^n$

$(\tilde{U}, \mathcal{B}(\tilde{U}), \nu) \longrightarrow$ probability space

$\mathcal{B}(\tilde{U})$: Borel set on U

$\boxed{\nu = \mathbb{P} \circ \xi^{-1}}$

For any open set O in $\mathbb{R}^n$ consider

$$S^{-1}(O) = \{ \xi \in U : S(\xi) \cap O \neq \emptyset \}$$

If for any open set $O \subset \mathbb{R}^n$, $S^{-1}(O) \in \mathcal{B}(U)$, then S is called "measurable".

The set $S(\xi)$ is called a random set if S is measurable

The expectation of a random set or a measurable multifunction

$$E_{\mathbb{P}}[S(\xi)] = \left\{ E_{\mathbb{P}}(v(\xi)) : v : \mathbb{R}^m \longrightarrow \mathbb{R}^n,\ v(\xi) \in S(\xi)\ \forall \xi \in U, \text{ and } v(\xi) \text{ integrable} \right\}$$

If $f : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ be smooth at $\bar{x}$,

If for any $\xi \in U$ define $S(\xi) = \{ \nabla_x f(\bar{x}, \xi) \}$

Then

$$E[S(\xi)] = \left\{ E_{\mathbb{P}} \{ \nabla_x f(\bar{x}, \xi) \} \right\}$$

Note $E[v(\xi)]$ depends on the probability distribution of ξ

$$\boxed{E[v(\xi)] = \int_{\Omega} v \circ \xi(\omega)\, d\mathbb{P} = \int_{\mathbb{R}^m} v(\xi)\, d\nu}$$

$v : U \longrightarrow \mathbb{R}^n$ denotes the random variable in the second expressio

Given a random vector $\xi : \Omega \to \mathbb{R}^m$, we define for the case where $f(\cdot, \xi)$ is convex for any $\xi \in \mathbb{R}^m$, the subdifferential of the expectation functional ϕ is given as

$$\partial \phi(x) = E_{\mathbb{P}}\left[\partial_x f(x, \xi)\right].$$

Is this a random set

Example (Computing the Subdifferential of expectation functional)
Source: (Royset and Wets : An Optimization Primer)

Beware of this Word

$$f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$$

with $f(x, \xi) = \max\{x - \xi, 0\}$.

$f(\cdot, \xi)$ is a non-smooth convex function.

Let ξ be a random variable with $U = \{-1, 0, 2\}$.

The probability masses are

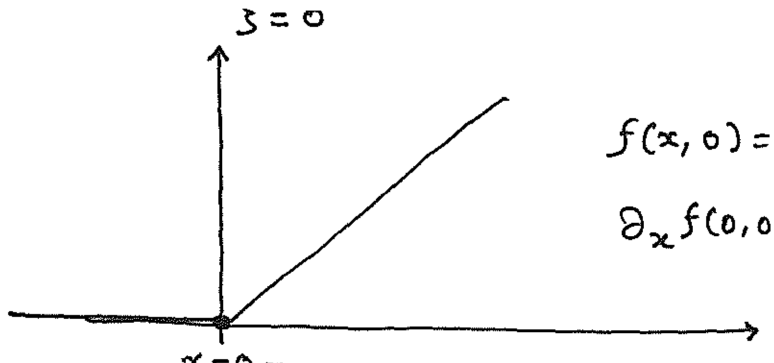
$$p_{-1} = 0.2, \quad p_0 = 0.5 \quad p_2 = 0.3$$

$$E_{\mathbb{P}}[f(x, \xi)] = \phi(x) = p_{-1} f(x, -1) + p_0 f(x, 0) + p_2 f(x, 2)$$

$$\phi(x) = 0.2 f(x, -1) + 0.5 f(x, 0) + 0.3 f(x, 2).$$

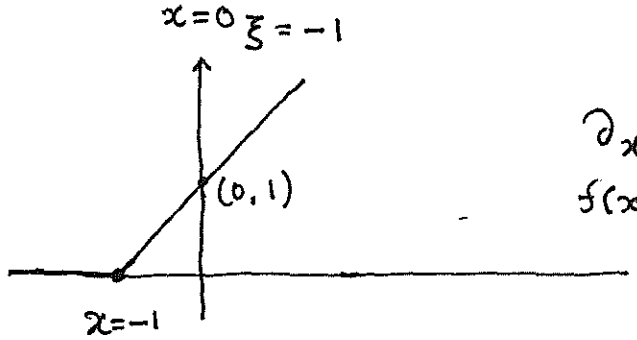
We will compute $E\left[\partial_x f(x, \xi)\right]$. We will consider the specific case of $x = 0$. Then

$$f(0, -1) = 1, \quad f(0, 0) = 0, \quad f(0, 2) = 0$$



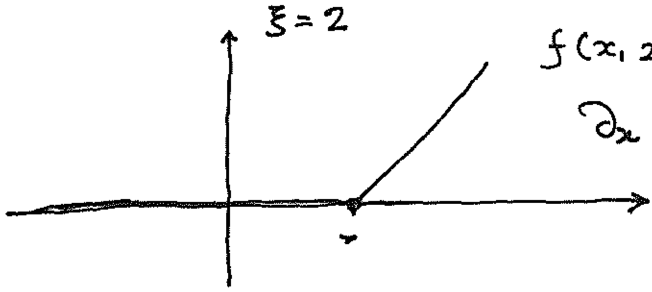
$f(x, 0) = \max\{x, 0\}$

$\partial_x f(0,0) = [0, 1]$



$\partial_x f(0,-1) = \{1\}$

$f(x, -1) = \max\{x+1, 0\}$



$f(x, 2) = \max\{x-2, 0\}$

$\partial_x f(0, 2) = \{0\}$

We have to find $v(\xi) \in \partial_x f(0, \xi)$

$v(-1) \in \partial_x f(0,-1) = \{1\}, \quad v(0) \in \partial_x f(0,0) = [0,1]$

$v(2) \in \partial_x f(0,2) = \{0\}$

$$\Rightarrow \mathbb{E}\left[v(\xi)\right] = 0.2\, v(-1) + 0.5\, v(0) + 0.3\, v(2)$$
$$\in [0.2, 0.7] = \mathbb{E}\left[\partial_x f(0, \xi)\right]$$
$$\partial \phi(0) = [0.2, 0.7].$$

Theoretical justification for

$$\partial \phi(x) = \mathbb{E}_{\mathbb{P}}\left[\partial_x f(x, \xi)\right].$$

- $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a measurable multi-function

Then if S is convex-valued then $E_{\mathbb{P}}[S(\xi)]$ is a convex set.

<u>Proof:</u> Let $\xi \in U = \text{Range}\, \xi$ and let $s^0, s^1 \in E_{\mathbb{P}}[S(\xi)]$

$\exists$ v^0 & v^1 : $\mathbb{R}^m \to \mathbb{R}^n$ with v^0, v^1 integrable

$v^0(\xi)$ & $v^1(\xi) \in S(\xi)$ $\forall\, \xi \in U$, such that

$$s^0 = E_{\mathbb{P}}[v^0(\xi)]$$
$$s^1 = E_{\mathbb{P}}[v^1(\xi)]$$

Now as S is <u>convex-valued</u>; then for any $\lambda \in [0,1]$

$$v^\lambda(\xi) = (1-\lambda)v^0(\xi) + \lambda v^1(\xi) \in S(\xi), \forall \xi$$

v^λ is integrable as

$$E_\nu[v^\lambda] = E_{\mathbb{P}}[v^\lambda(\xi)] = (1-\lambda)E_{\mathbb{P}}[v^0(\xi)] + \lambda E_{\mathbb{P}}[v^1(\xi)] < \infty$$

Further

$$E_{\mathbb{P}}[v^\lambda(\xi)] = (1-\lambda)s^0 + \lambda s^1$$

Since $E_{\mathbb{P}}[v^\lambda(\xi)] \in E_{\mathbb{P}}[S(\xi)]$

Thus $(1-\lambda)s^0 + \cancel{(1-\lambda)s}\,\lambda s^1 \in E_{\mathbb{P}}[S(\xi)]$

Voila!! <u>We are done!!</u>

- Let ξ be a random vector and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

If $f(x, \xi)$ is integrable for all x and ~~$f(\xi, \cdot)$~~ $f(\cdot, \xi)$ is convex for all $\xi \in U$, then $\partial_x f(x, \xi)$ is a random set for all $x \in \mathbb{R}^n$. Further

$$\partial \Phi(x) = E_{\mathbb{P}}\left[\partial_x f(x, \xi)\right]$$

If x^* be a minimizer of ϕ over $\mathbb{R}^n$, then

$$0 \in \partial \phi(x^*)$$

$$\Rightarrow \exists\, v : \mathbb{R}^m \to \mathbb{R}^n, \quad v \text{ is integrable and}$$

$$v(\xi) \in \partial_x f(x^*, \xi), \quad \forall \xi \in U.$$

The converse also holds

End of Part-1

Part-II : Basic models of stochastic optimization and measures of risk