Advanced Topics in OR

Lecture Notes 3

Stochastic Order Relations

Zhao Xiaobo
Department of IE
Tsinghua University
Beijing 100084, China
Tel. 010-62784898
Email. xbzhao@tsinghua.edu.cn

Likelihood Ratio Ordering

X and Y: nonnegative, densities f and g

$$X \geq_{LR} Y \qquad \text{if} \qquad \frac{f(x)}{g(x)} \leq \frac{f(y)}{g(y)} \qquad \text{for all } x \leq y$$

Hence, $X \geq_{LR} Y$ if the ratio of their respective densities, $f(x)/g(x)$, is nondecreasing in x.

Proposition 9.4.1: If $X \geq_{LR} Y$, then $\lambda_X(t) \leq \lambda_Y(t)$ for all $t \geq 0$.

Proof: Since $X \geq_{LR} Y$, it follows that, for $x \geq t$,

$$f(x) \geq g(x)\, f(t)/g(t)$$

Likelihood Ratio Ordering

$$\Rightarrow \lambda_X(t) = \frac{f(t)}{\int_t^\infty f(x)\,dx} \leq \frac{f(t)}{\int_t^\infty g(x) f(t)/g(t)\,dx}$$

$$= \frac{g(t)}{\int_t^\infty g(x)\,dx} = \lambda_Y(t)$$

Example 9.4(A): $X \sim$ exponential λ, $Y \sim$ exponential μ

$$\Rightarrow \quad \frac{f(x)}{g(x)} = \frac{\lambda}{\mu} e^{(\mu-\lambda)x}$$

So, $X \geq_{LR} Y$ when $\lambda \leq \mu$.

Likelihood Ratio Ordering

Example 9.4(B): ***A statistical inference problem***

Inference about unknown distribution of a given random variable

X: a continuous random variable possibly having density functions f or g.

Based on the observed value of X, we must decide on either f or g.

A function $\varphi(x)$ takes either value 0 or 1: a decision rule

$$\varphi(x) = 0, \text{ decide on } f$$

$$\varphi(x) = 1, \text{ decide on } g$$

Likelihood Ratio Ordering

Note that the probability of rejecting f when it is true density

$$\int_{x:\varphi(x)=1} f(x)\,dx = \int f(x)\varphi(x)\,dx$$

An approach: fix a constant α, $0 \le \alpha \le 1$,

$$\int f(x)\varphi(x)\,dx \le \alpha$$

Among such rules, we choose the one that maximizes the probability of rejecting f when it is false.

$$\int_{x:\varphi(x)=1} g(x)\,dx = \int g(x)\varphi(x)\,dx$$

The optimal decision rule: Neyman-Pearson lemma

Likelihood Ratio Ordering

Neyman-Pearson Lemma: Among all decision rules φ, the one maximizes $\int g(x)\varphi(x)dx$ is φ^* given by

$$\varphi^*(x)=\begin{cases}0 & \text{if } f(x)/g(x)\geq c\\ 1 & \text{if } f(x)/g(x)< c\end{cases}$$

where c is chosen so that

$$\int f(x)\varphi^*(x)dx=\alpha$$

Proof: For any x

$$\left(\varphi^*(x)-\varphi(x)\right)\left(cg(x)-f(x)\right)\geq 0$$

The above follows: if $\varphi^*(x) = 1$, then both terms in the product are nonnegative, and if $\varphi^*(x) = 0$, then both are nonpositive

Likelihood Ratio Ordering

$$\Rightarrow \quad \int\left(\varphi^*(x)-\varphi(x)\right)\left(cg(x)-f(x)\right)dx \geq 0$$

$$\Rightarrow \quad c\left[\int\varphi^*(x)g(x)dx-\int\varphi(x)g(x)dx\right] \geq$$

$$\int\varphi^*(x)f(x)dx-\int\varphi(x)f(x)dx \geq 0$$

Suppose that *f* and *g* have a monotone likelihood ratio order: $f(x)/g(x)$ is nondecreasing in x.

The optimal decision rule $\varphi^*(x)=\begin{cases}0 & \text{if } x \geq k \\ 1 & \text{if } x < k\end{cases}$

where k is such that $\int_{-\infty}^{k} f(x)dx=\alpha$

Likelihood Ratio Ordering

Important applications in optimization theory

Proposition 9.4.2: Suppose that X and Y are independent with densities f and g, and $X \geq_{LR} Y$. If $h(x, y)$ is a real-valued function satisfying

$$h(x,y) \geq h(y,x) \quad \text{whenever } x \geq y,$$

then $h(X,Y) \geq_{st} h(Y,X)$

Proof: Let $U = \max(X, Y)$, $V = \min(X, Y)$. Conditional on $U = u$, $V = v$, $u \geq v$, the conditional distribution of $h(X, Y)$ is concentrated on the two points $h(u, v)$ and $h(v, u)$

Assigning the following probability to the larger value $h(u, v)$

Likelihood Ratio Ordering

$$\lambda_1 \equiv P\left\{X = \max\left(X, Y\right), Y = \min\left(X, Y\right) \middle| U = u, V = v\right\}$$

$$= \frac{f\left(u\right)g\left(v\right)}{f\left(u\right)g\left(v\right) + f\left(v\right)g\left(u\right)}$$

Conditional on $U = u$, $V = v$, $u \geq v$, $h(Y, X)$ is also concentrated on the two points $h(u, v)$ and $h(v, u)$

Assigning the probability to the value $h(v, u)$

$$\lambda_2 \equiv P\left\{Y = \max\left(X, Y\right), X = \min\left(X, Y\right) \middle| U = u, V = v\right\}$$

$$= \frac{f\left(v\right)g\left(u\right)}{f\left(u\right)g\left(v\right) + f\left(v\right)g\left(u\right)}$$

Likelihood Ratio Ordering

Since $u \geq v$

$$f(u)g(v) \geq f(v)g(u)$$

⟹ Conditional on $U = u$, $V = v$, $h(X, Y)$ is stochastically larger than $h(Y, X)$, i.e.,

$$P\{h(X,Y) \geq a | U = u, V = v\} \geq P\{h(Y,X) \geq a | U = u, V = v\}$$

Remark: The above does not necessarily hold when we only assume that $X \geq_{st} Y$.

A counter example: $2x + y \geq x + 2y$ whenever $x \geq y$

$$X = \begin{cases} 3 & \text{with probability } 0.2 \\ 9 & \text{with probability } 0.8 \end{cases} \qquad Y = \begin{cases} 1 & \text{with probability } 0.2 \\ 4 & \text{with probability } 0.8 \end{cases}$$

Likelihood Ratio Ordering

⇒ $X \geq_{st} Y$

But $P\{2X + Y \geq 11\} = 0.8$ and $P\{2Y + X \geq 11\} = 0.8 + 0.2 \times 0.8 = 0.96$.

Thus, $2X + Y$ is not stochastically larger than $2Y + X$

Proposition 9.4.2 has important applications in optimal scheduling problems.

n items, each has some measurable characteristic, e.g., processing time.

x_i: characteristic of item i

$h(x_{i1}, x_{i2}, \ldots, x_{in})$: return of order chosen $i_1, \ldots, i_n$

Likelihood Ratio Ordering

Let the characteristic of item *i* be a random variable

if $$X_1 \geq_{LR} X_2 \geq_{LR} \cdots \geq_{LR} X_n$$

and $$h(y_1, \cdots, y_{i-1}, y_i, y_{i+1}, \cdots, y_n) \geq h(y_1, \cdots, y_i, y_{i-1}, y_{i+1}, \cdots, y_n)$$

whenever $y_i > y_{i-1}$, then it follows from Proposition 9.4.2 that the ordering 1, 2, …, n (n, $n-1$, …, 1) stochastically maximizes (minimizes) the return.

Consider any ordering that does not start with item 1 – say (i_1, i_2, 1, …, i_{n-1}).

By conditioning on the values $X_{i_1}, X_{i_3}, \cdots, X_{i_{n-1}}$

⇒ Proposition 9.4.2 shows that the ordering (i_1, 1, i_2, …, i_{n-1}) leads to a stochastically larger return.

Likelihood Ratio Ordering

Continuing with such interchanges leads to the conclusion that 1, 2, …, n stochastically maximizes the return.

A similar argument shows that n, $n - 1$, …, 1 stochastically minimizes return.

X: density f

log(f(x)) is concave → increasing likelihood ratio

log(f(x)) is convex → decreasing likelihood ratio

random variable $c + X$: density $f(x - c)$

$$c_1 + X \geq_{LR} c_2 + X \quad \text{for all } c_1 \geq c_2 \Leftrightarrow \frac{f(x - c_1)}{f(x - c_2)} \uparrow x \quad \text{for all } c_1 \geq c_2$$

Likelihood Ratio Ordering

$$\Leftrightarrow \quad \log f(x-c_1) - \log f(x-c_2) \uparrow x \quad \text{for all } c_1 \geq c_2$$

$$\Leftrightarrow \quad \log f(x) \text{ is concave}$$

Hence, X has increasing likelihood ratio if $c + X$ increases in likelihood ratio as c increases

A second interpretation: X_t, the remaining life from t onward, having lifetime X, which has reached the age of t

$$\bar{F}_t(a) \equiv P\{X_t > a\} = \frac{\bar{F}(t+a)}{\bar{F}(t)}$$

The density $\quad f_t(a) = \dfrac{f(t+a)}{\bar{F}(t)}$

Likelihood Ratio Ordering

Hence

$$X_s \geq_{LR} X_t \quad \text{for all } s \leq t \quad \Leftrightarrow \quad \frac{f(s+a)}{f(t+a)} \uparrow a \quad \text{for all } s \leq t$$

$$\Leftrightarrow \quad \log f(x) \text{ is concave}$$

Therefore, X has increasing likelihood ratio if X_s decreases in likelihood ratio as s increases

Proposition 9.4.3: If X has increasing likelihood ratio, then X is IFR. If X has decreasing likelihood ratio, then X is DFR.

Proof: $X_s \geq_{LR} X_t \quad \Rightarrow \quad \lambda_{X_s} \leq \lambda_{X_t}$

$$\Rightarrow \quad X_s \geq_{st} X_t$$

Likelihood Ratio Ordering

Remarks:

1) A density function f such that $\log f(x)$ is concave is called a Polya frequency of order 2.

2) For discrete random variables, $X \geq_{LR} Y$ if $P\{X = x\}/P\{Y = x\}$ increases in x.

Stochastically More Variable

h is convex if for all $0 < \lambda < 1$, x_1, x_2,

$$h\left(\lambda x_1 + (1-\lambda)x_2\right) \leq \lambda h(x_1) + (1-\lambda)h(x_2)$$

X is more variable than Y, write $X \geq_V Y$, if

$$E\left[h(X)\right] \geq E\left[h(Y)\right] \quad \text{for all increasing, convex } h$$

If X and Y have distributions F and G, then $F \geq_V G$ when the above holds

Proposition 9.5.1: If X and Y are nonnegative variables with distributions F and G, then $X \geq_V Y$, iff

$$\int_a^{\infty} \overline{F}(x)\,dx \geq \int_a^{\infty} \overline{G}(x)\,dx \quad \text{for all } a \geq 0$$

Stochastically More Variable

Proof: Let h_a be defined by

$$h_a = (x-a)^+ = \begin{cases} 0 & x \le a \\ x-a & x > a \end{cases}$$

Since h_a is increasing and convex, we have for $X \ge_V Y$,

$$E\left[h_a(X)\right] \ge E\left[h_a(Y)\right]$$

But

$$E\left[h_a(X)\right] = \int_0^\infty P\left\{(X-a)^+ > x\right\} dx$$

$$= \int_0^\infty P\{X > a + x\} dx$$

$$= \int_a^\infty \bar{F}(y) dy$$

Stochastically More Variable

Similarly, we have

$$E\left[h_a(Y)\right]=\int_a^{\infty}\bar{G}(y)dy$$

Go the other way: assume *h* is twice differentiable

Since *h* convex means $h'' \geq 0$, we have

$$\int_0^{\infty}h''(a)\int_a^{\infty}\bar{F}(x)dxda \geq \int_0^{\infty}h''(a)\int_a^{\infty}\bar{G}(x)dxda$$

The left-hand side

$$\int_0^{\infty}h''(a)\int_a^{\infty}\bar{F}(x)dxda = \int_0^{\infty}\int_0^{x}h''(a)da\bar{F}(x)dx$$

$$=\int_0^{\infty}h'(x)\bar{F}(x)dx - h'(0)E[X]$$

Stochastically More Variable

$$= \int_0^\infty h'(x) \int_x^\infty dF(y) dx - h'(0) E[X]$$

$$= \int_0^\infty \int_0^y h'(x) dx dF(y) - h'(0) E[X]$$

$$= \int_0^\infty h(y) dF(y) - h(0) - h'(0) E[X]$$

$$= E[h(X)] - h(0) - h'(0) E[X]$$

A similar identity is valid for G. We see

$$E[h(X)] - E[h(Y)] \geq h'(0)\left(E[X] - E[Y]\right)$$

The right-hand side is nonnegative since $E[X] \geq E[Y]$ by setting $a = 0$

Stochastically More Variable

Corollary 9.5.2: If X and Y are nonnegative variables such that $E[X] = E[Y]$, then $X \geq_V Y$, iff

$$E\left[h(X)\right] \geq E\left[h(Y)\right] \quad \text{for all convex } h$$

Proof: Let h be convex and suppose that $X \geq_V Y$. Then as $E[X] = E[Y]$, from the convex h, we have

$$E\left[h(X)\right] \geq E\left[h(Y)\right]$$

For two nonnegative random variables having the same mean, we have that $X \geq_V Y$ if $E[h(X)] \geq E[h(Y)]$ for all convex functions h.

Stochastically More Variable

$X \geq_V Y$ means that X has more variability than Y.

Intuitively, X is more variable than Y if it gives more weight to the extreme values, and one way of guaranteeing this is to require that $E[h(X)] \geq E[h(Y)]$ whenever h is convex.

For instance, since $E[X] = E[Y]$ and since $h(x) = x^2$ is convex, we would have that $\mathrm{Var}(X) \geq \mathrm{Var}(Y)$.

Corollary 9.5.3: If X and Y are nonnegative variables with $E[X] = E[Y]$, then $X \geq_V Y$ implies that $-X \geq_V -Y$

Proof: Let h be an increasing and convex function. We have

$E\left[h\left(-X\right)\right] \geq E\left[h\left(-Y\right)\right]$ because $f(x) = h(-x)$ is convex

Stochastically More Variable

Proposition 9.5.4: If $X_1, \ldots, X_n$ are independent and $Y_1, \ldots, Y_n$ are independent, and $X_i \geq_V Y_i$, $i = 1, \ldots, n$, then

$$g(X_1, \cdots, X_n) \geq_V g(Y_1, \cdots, Y_n)$$

for all increasing convex functions g that are convex in each argument.

Proof: Start by assuming that the set of $2n$ random variables is independent. The proof is by induction on n.

When $n = 1$, we must show

$$E\left[h\left(g\left(X_1\right)\right)\right] \geq E\left[h\left(g\left(Y_1\right)\right)\right]$$

when g and h are increasing and convex and $X_1 \geq_V Y_1$

Stochastically More Variable

Since

$$\frac{d}{dx}h\left(g\left(x\right)\right)=h^{'}\left(g\left(x\right)\right)g^{'}\left(x\right)\geq 0$$

$$\frac{d^{2}}{dx^{2}}h\left(g\left(x\right)\right)=h^{"}\left(g\left(x\right)\right)\left(g^{'}\left(x\right)\right)^{2}+h^{'}\left(g\left(x\right)\right)g^{"}\left(x\right)\geq 0$$

Thus, $h(g(x))$ is increasing and convex, so the result follows.

Assume the result for vectors of size n – 1.

$$E\left[h\left(g\left(X_1,X_2,\cdots,X_n\right)\right)\middle|X_1=x\right]=E\left[h\left(g\left(x,X_2,\cdots,X_n\right)\right)\middle|X_1=x\right]$$

$$=E\left[h\left(g\left(x,X_2,\cdots,X_n\right)\right)\right]$$

$$\geq E\left[h\left(g\left(x,Y_2,\cdots,Y_n\right)\right)\right]$$

Stochastically More Variable

$$= E\left[h\left(g\left(X_1, Y_2, \cdots, Y_n\right)\right) \middle| X_1 = x\right]$$

Taking expectations gives that

$$E\left[h\left(g\left(X_1, X_2, \cdots, X_n\right)\right)\right] \geq E\left[h\left(g\left(X_1, Y_2, \cdots, Y_n\right)\right)\right]$$

Using the result for n = 1, we can show

$$E\left[h\left(g\left(X_1, Y_2, \cdots, Y_n\right)\right)\right] \geq E\left[h\left(g\left(Y_1, Y_2, \cdots, Y_n\right)\right)\right]$$

which prove the result.

2n random variables being independent do not affect the distributions of $g(X_1, \ldots, X_n)$ and $g(Y_1, \ldots, Y_n)$.

The result remains true under the weaker hypothesis that the two sets of n random variables are independent.