$$\begin{aligned}\int_a^b u(W)\, f(W)\, dW &= [\, u(W)\, F(W)\,]_a^b - \int_a^b u'(W)\, F(W)\, dW \qquad \text{integrating by parts}\\ &= u(b) * 1 - u(a) * 0 - \int_a^b u'(W)\, F(W)\, dW\\ &= u(b) - \int_a^b u'(W)\, F(W)\, dW\end{aligned}$$

Therefore comparing expected utility under two different distributions (if they have different "supports" $[a, b]$, take the biggest and define the other density outside its support to be 0):

$$\begin{aligned}\int_a^b u(W)\, f_1(W)\, dW - \int_a^b u(W)\, f_2(W)\, dW &= -\int_a^b u'(W)\, F_1(W)\, dW + \int_a^b u'(W)\, F_2(W)\, dW\\ &= \int_a^b u'(W)\, [\, F_2(W) - F_1(W)\,]\, dW\,.\end{aligned}$$

as a theorem:

Definition 1: The distribution F_1 is first-order stochastic dominant over F_2 if and only if $F_1(W) < F_2(W)$ for all $W \in (a, b)$.

Theorem 1: Every expected-utility maximizer with an increasing utility function of wealth prefers the lottery L^1 with distribution F_1 to L^2 with distribution F_2 if and only if F_1 is FOSD over F_2.

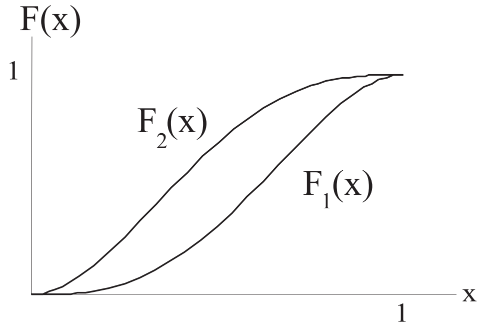

Figure 1: FOSD: CDF comparison

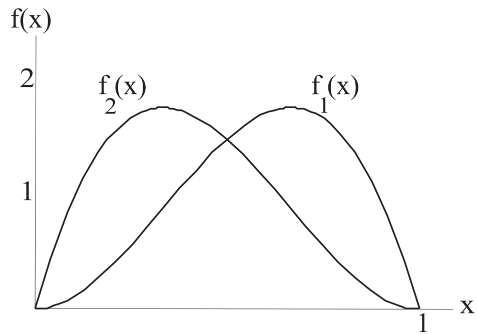

Figure 2: FOSD: Density comparison

$$
\begin{aligned}
\int_a^b u(W)\, f(W)\, dW &= [\, u(W)\, F(W)\,]_a^b - \int_a^b u'(W)\, F(W)\, dW \qquad \text{integrating by parts} \\
&= u(b) * 1 - u(a) * 0 - \int_a^b u'(W)\, F(W)\, dW \\
&= u(b) - \int_a^b u'(W)\, F(W)\, dW \\
&= u(b) - [\, u'(W)\, S(W)\,]_a^b + \int_a^b u''(W)\, S(W)\, dW \quad \text{int. by parts again} \\
&= u(b) - u'(b)\ S(b) + \int_a^b u''(W)\, S(W)\, dW
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\int_a^b u(W)\, f_1(W)\, dW - \int_a^b u(W)\, f_2(W)\, dW \\
= \ & u'(b)\, [S_2(b) - S_1(b)] + \int_a^b u''(W)\, [S_1(W) - S_2(W)]\, dW
\end{aligned}
$$

$$\begin{aligned}
S(b) &= \int_a^b F(W)\, dW = \int_a^b F(W) * 1\, dW \\
&= [\, F(W)\, W\,]_a^b - \int_a^b f(W)\, W\, dW \qquad \text{int. by parts} \\
&= F(b)\, b - F(a)\, a - \int_a^b f(W)\, W\, dW \\
&= b - E[W]
\end{aligned}$$

Therefore, if the two distributions have $S_1(b) = S_2(b)$, then the expected values of W under the two distributions should also be equal, which we can write as $E_1(W) = E_2(W)$. In other words, the mean wealth should be the same under the two lotteries. This is a very natural condition to require when we want preference between them to depend only on the attitudes toward risk.

So we are left with

$$\int_a^b u(W)\, f_1(W)\, dW - \int_a^b u(W)\, f_2(W)\, dW = \int_a^b u''(W)\, [S_1(W) - S_2(W)]\, dW\,.$$

Proceeding exactly as the case of first-order stochastic dominance, we see that the right-hand

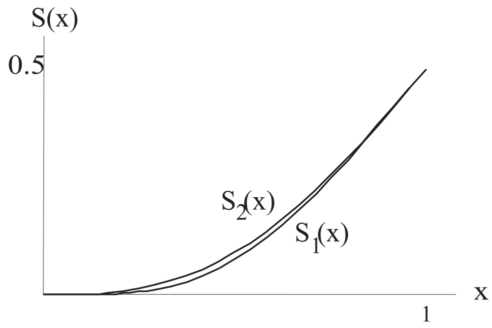

Figure 3: SOSD: S-function compariison

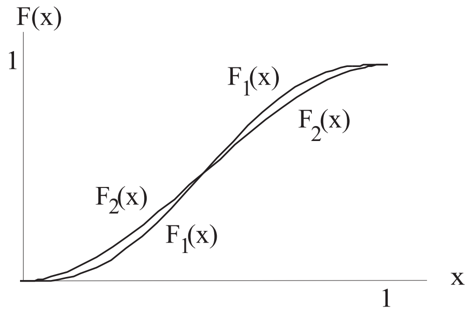

Figure 4: SOSD: CDF comparison

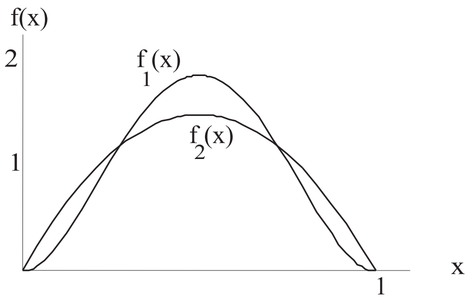

Figure 5: SOSD: density comparison

Definition 2(a): Let W_i denote the random variable following distribution F_i, for $i = 1$, 2, and $E_1(W_1) = E_2(W_2)$. Then F_1 is said to be SOSD over F_2 if there exists a random variable z with zero expectation conditional on any given value of W_1, such that w_2 has the same distribution as $w_1 + z$, or in other words, w_2 equals W_1 plus some added pure uncertainty or "noise."

Definition 2(b): Of two distributions yielding equal expected values, F_1 is said to be SOSD over F_2 if it is possible to get from F_1 to F_2 by a sequence of operations which shift pairs of probability weights on either side of the mean farther away, while leaving the mean unchanged.

The two lotteries, defined by their vectors of wealth consequences and probabilities, are:

$L^1 = (0, 2, 4, 6;\ 1/4, 1/4, 1/4, 1/4)$
$L^2 = (0.1, 3, 5.9;\ 1/3, 1/3, 1/3)$

It is easy to calculate that the means and variances are

L^1: Mean $= 3$, Variance $= 5.000$
L^2: Mean $= 3$, Variance $= 5.607$

Take the utility function

$$u(W) = \begin{cases} 2W & \text{if } W \leq 3 \\ 3 + W & \text{if } W > 3 \end{cases}$$

Figure 6 shows the two lotteries and Figure 7 shows the utility function:

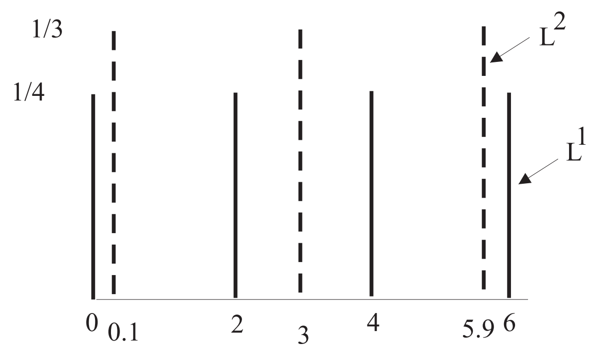

Figure 6: Lotteries in example for SOSD $\neq$ lower variance

This gives expected utilities:

$EU(L^1) = 5.000, \qquad EU(L^2) = 5.033$

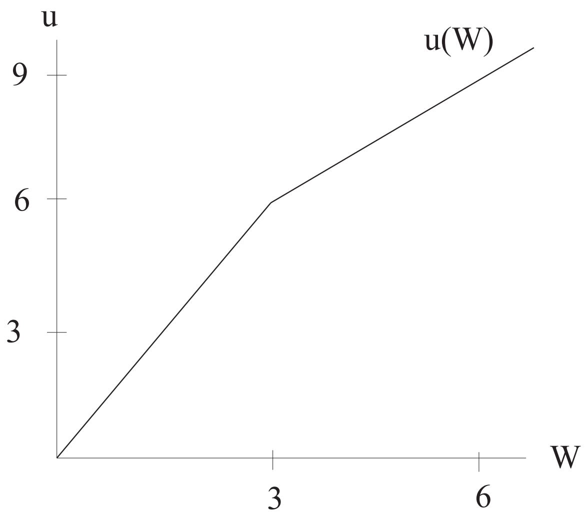

Figure 7: Utility in example for SOSD *neq* low variance

so the person prefers L^2 despite the fact that it has a higher variance than L^1 and the same mean.

And how does this relate to the definition of SOSD? For that, we need to compare the super-cumulative functions for the two distributions. A little work shows yields Figure 6. I have shown S_1 thicker, and have shown the points 0.1 and 5.9 out of scale for clarity of appearance. We see that the two super-cumulatives S_1 and S_2 cross, so F_1 is not SOSD over F_2.

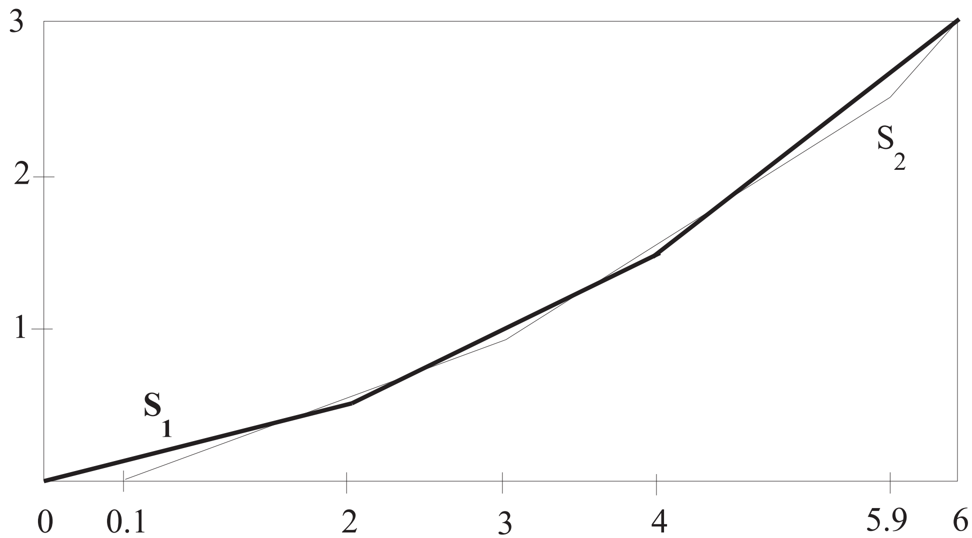

Figure 8: SOSD vs. variance comparison

Another way to look at this is that SOSD is a kind of comprehensive requirement that