Last updated: October 16, 2017

PhD Micro (Part 1)
Risk

Motaz Al-Chanati

October 10, 2017

1 Risk Aversion

1.1 Framework

In the previous section, I said it is easiest to think about the prizes from lotteries as being from money. That's going to be especially true here, but we will only consider lotteries over non-negative dollar amounts: $X \subset \mathbb{R}_+$. A big difference here is that instead of having a discrete set of prizes and a discrete probability vector, we're going to consider lotteries as being represented by a *continuous* probability distribution. In particular, the prizes x will come from a cumulative distribution function $F : X \to [0, 1]$ that has finite support.[1] Just as we had infinitely many valid probability vectors p, any distribution function F with support in X can represent a possible lottery.

Let's go over some useful properties of distribution functions:

- The cumulative distribution function (CDF) $F(x)$ is non-decreasing and represents the probability of getting a prize less than or equal to x, i.e. $F(x) = P(X \leq x)$.
- The probability density function (PDF) $f(x)$ represents how likely you are to get a prize equal to x (more precisely, in a infinitesimally small neighborhood around x).[2]
- $F(x) = \int_{-\infty}^{x} f(t)dt$ and $\frac{dF(x)}{dx} = f(x) \iff dF(x) = f(x)dx$ (since we are working with continuous lotteries)
- $\int_{x\in X} dF(x) = 1$ (support of F is in X)
- The expected value of lottery with CDF F is $\mu_F = \int_{x\in X} xdF(x)$

Just like before, we will have preferences (and utility representation) over these lotteries. We are going to still assume preferences are rational, continuous and satisfy the independence axiom. This means they have vNM (expected utility) form. The continuous analogue of "summing over" utilities with the sum operator Σ is simply integrating $\int$. As in the last section, utility functions over lottery are $U(\cdot)$ and utility functions over prizes are $u(\cdot)$ (the Bernoulli utility). So the expected utility form here means:

$$U(F) = \int_{x\in X} u(x)dF(x) = \int_{x\in X} u(x)f(x)dx$$

[1] Before we had a finite number of prizes, so this is just the continuous analogue of that same principle

[2] Following usual convention, we will have uppercase letters representing CDFs and the corresponding lowercase letter representing its PDF

Applying this to utility representation, we get for any two lotteries (CDFs) F and G:

$$F \succsim G \iff U(F) \geq U(G) \iff \int_{x \in X} u(x)dF(x) \geq \int_{x \in X} u(x)dG(x)$$

We will usually assume that $u(\cdot)$ is a bounded, continuous, and strictly increasing increasing function. The Bernoulli utility function $u(x)$ will again be defined as the utility from the corresponding degenerate lottery δ_x: $u(x) = U(\delta_x)$

1.2 Key Concepts

If you remember in the last section, I said that $U(F)$ represented the expectation over utilities, which in general is not the utility of the expectation. The comparison between the two gives us the following terms:

Risk Attitudes
An agent with preferences $\succsim$ and Bernoulli utility function $u(\cdot)$ has the following risk attitude, if **for all** lotteries F:

- *Risk averse:* $\delta_{\mu_F} \succsim F$, equivalently, $u(\mu_F) \geq U(F)$.
- *Risk neutral:* $\delta_{\mu_F} \sim F$, equivalently, $u(\mu_F) = U(F)$.
- *Risk loving:* $F \succsim \delta_{\mu_F}$, equivalently, $u(\mu_F) \leq U(F)$.

So in other words, a risk averse agent prefers getting the mean of the lottery for sure rather than enter the lottery. If we write this out in full vNM form we get:

$$\begin{aligned} & u(\mu_F) \geq U(F) \\ \iff & u\left(\int_{x \in X} x dF(x)\right) \geq \int_{x \in X} u(x)dF(x) \end{aligned}$$

But for this last inequality to hold, we need to have $u(\cdot)$ be concave because this is just Jensen's inequality.[3] So we can think of risk aversion equivalently as being concavity of the Bernoulli function. By a similar logic, risk neutrality means that $u(\cdot)$ is linear and risk loving means that $u(\cdot)$ is convex.

The next natural question is to think if getting μ_F for sure does not give the same utility as F, then is there some other degenerate lottery that would make the agent indifferent? We call this concept the *certainty equivalent* of a lottery.

Risk Concepts
An agent with preferences $\succsim$ and Bernoulli utility function $u(\cdot)$, the following concepts are defined **for all** lotteries F:

- *Certainty equivalent:* A real number $c(F, u) \in \mathbb{R}$ such that $\delta_{c(F,u)} \sim F$, equivalently, $u(c(F, u)) = U(F)$.
- *Risk premium:* The difference between expected value and the certainty equivalent of a lottery: $\rho(F, u) = \mu_F - c(F, u)$

[3] Jensen's inequality says that for a random variable X and a concave function $f(\cdot)$, $f(E[X]) \geq E[f(x)]$ (and $\leq$ if f is convex)

For a risk averse agent, since $u(\mu_F) \geq U(F)$ and $u(\cdot)$ is increasing, we need a number c smaller than μ_F to make $u(c) = U(F)$. Therefore, we must have $c(F, u) \leq \mu_F$ and thus $\rho(F, u) \geq 0$. Another way to express these concepts is the following:

$$u(c(F,u)) = U(F)$$
$$u(\mu_F - \rho(F,u)) = U(F)$$

2 Measuring Risk Aversion

2.1 Indices

Right now all we have is just a way of saying whether someone is risk averse or now. We would like to have some sort of measure of *how much* risk aversion a person exhibits. There are going to be three things we're going to look for in this measure:

1. It should be invariant to an increasing affine transformation of $u(\cdot)$.
   - We know that with vNM utility form, we can replace the Bernoulli utility function with $v(x) = au(x) + b$ for $a > 0$ and the new utility function still represents the same preference. Clearly we wouldn't our risk aversion measure to change just because we are using a different utility representation
2. It should involve $u''(\cdot)$
   - We know that risk aversion is equivalent to concavity of the Bernoulli utility function. The second derivative measures how concave a function is ($u'' < 0$ means u is concave)
3. It must be a local measure
   - The concavity of $u(x)$ can be changing at different values of x. So how risk averse you are should depend on x (i.e. you might be more or less risk averse depending on how much is at stake)

Let's start by considering now a 'local' lottery around a prize x (call it $\tilde{x}$). By local, we mean that you could win either $x + \varepsilon$ with 0.5 probability or $x - \varepsilon$ with 0.5 probability (for some small $\varepsilon > 0$). Using the definitions above:

- Mean of the lottery: $\frac{1}{2}(x + \varepsilon) + \frac{1}{2}(x - \varepsilon) = x$
- Expected utility: $U(\tilde{x}) = \frac{1}{2}[u(x + \varepsilon) + u(x - \varepsilon)]$
- Risk premium: $\rho_{\tilde{x}}$ such that $u(\mu_{\tilde{x}} - \rho_{\tilde{x}}) = u(x - \rho_{\tilde{x}}) = \frac{1}{2}[u(x + \varepsilon) + u(x - \varepsilon)]$

Using the risk premium equation, we can take first order Taylor expansions of $u(\cdot)$. Recall that the Taylor expansion of $f(y)$ around y^* is $f(y) = f(y^*) + f'(y^*)(y - y^*) + \frac{1}{2}f''(y^*)(y - y^*)^2 + \ldots$ (other small terms that I will drop for simplicity)[4] In our case, we have $y = x - \rho_{\tilde{x}}$ and $y^* = x$, so for the LHS:

$$u(x - \rho_{\tilde{x}}) \approx u(x) + u'(x)(-\rho_{\tilde{x}})$$

[4] In this next part, I'm not going to be very rigorous but if you want to see this done more formally, look at Chapter 6 of the Kreps Student Guide, Ex 6.3: https://sites.google.com/a/stanford.edu/microfoundations1/

And for the RHS, we do it twice for each $u(\cdot)$. One with $y = x + \varepsilon$ and $y^* = x$ and the other with $y = x - \varepsilon$ and $y^* = x$

$$\begin{aligned}\frac{1}{2}[u(x+\varepsilon) + u(x-\varepsilon)] &\approx \frac{1}{2}[u(x) + u'(x)\varepsilon] + \frac{1}{2}[u(x) + u'(x)(-\varepsilon)] \\ &\approx u(x)\end{aligned}$$

Putting this together, the $u(x)$'s cancel out and since $u' > 0$ (it is an increasing function), then we must have that $\rho_{\tilde{x}} \approx 0$. This gives us the idea of *first order risk neutrality.* In other words, if you play a very small local lottery, all agents look approximately risk neutral. This makes intuitive sense too; the first order Taylor approximation is just a (local) linearization of the utility function and we know that risk neutral agents have linear utility curves.

Ok, so now let's go one step further and do a second order expansion on the RHS because we know that u'' has to appear somewhere:

$$\begin{aligned}\frac{1}{2}[u(x+\varepsilon) + u(x-\varepsilon)] \approx &\frac{1}{2}\left[u(x) + u'(x)\varepsilon + \frac{1}{2}u''(x)\varepsilon^2\right] \\ &+ \frac{1}{2}\left[u(x) + u'(x)(-\varepsilon) + \frac{1}{2}u''(x)(-\varepsilon)^2\right] \\ \approx &u(x) + \frac{1}{2}u''(x)\varepsilon^2\end{aligned}$$

Putting this together:

$$\begin{aligned}u(x) + u'(x)(-\rho_{\tilde{x}}) &\approx u(x) + \frac{1}{2}u''(x)\varepsilon^2 \\ \rho_{\tilde{x}} &\approx \frac{\varepsilon^2}{2} \cdot \left(-\frac{u''(x)}{u'(x)}\right)\end{aligned}$$

Since our ε lottery is mean zero, you can interpret ε^2 as the variance of the local lottery. So if we change ε this part changes, but the part inside the brackets stays the same. So we can use it as a measure of risk aversion (for local lotteries around x). Clearly it uses $u''(x)$ and is local so let's check it is invariant to an affine transformation. Let $v(x) = au(x) + b$, so $v''(x) = au''(x)$ and $v'(x) = au'(x)$. The a's will cancel and the ratio will still be the same too.

Now that we've done all this, let's introduce our two main measures of risk aversion

Measures of Risk Aversion
For an agent with twice differentiable Bernoulli utility function $u(\cdot)$, the following are measures for the agent's degree of risk aversion

- *Coefficient/Index of absolute risk aversion:*

$$I_A(x, u) = -\frac{u''(x)}{u'(x)}$$

- *Coefficient/Index of relative risk aversion:*

$$I_R(x, u) = -x\frac{u''(x)}{u'(x)}$$

The coefficient of absolute risk aversion is also often called the Arrow-Pratt measure. The intuition is actually quite easy. As we've already said, u'' measures the concavity and therefore the risk aversion of an agent. But since that is not unique to affine transformation, dividing it by u' essentially normalizes it (and the negative sign helps give it the interpretation that a larger number is more risk averse).

In our derivation for a measure of absolute risk aversion, we considered a local lottery where you could gain or lose ε. This is an absolute dollar amount and has nothing to do with your current wealth level x. If we instead made the lottery a *percentage* of your wealth, i.e. you could gain or lose αx. For a given level of x, let's define a new Bernoulli utility function $\tilde{u}(\cdot)$ such that $\tilde{u}(\alpha) = u(\alpha x)$. Since we are looking for local lotteries around the current wealth level, we want to look for risk aversion around $\alpha = 1$. But if we look at a lottery $\tilde{u}(1+\varepsilon)$ and $\tilde{u}(1-\varepsilon)$, that will just give us the same absolute risk aversion index, but using the new utility function $\tilde{u}$:

$$-\frac{\tilde{u}''(1)}{\tilde{u}'(1)} = -\frac{\left.\frac{\partial^2 u(\alpha x)}{\partial \alpha^2}\right|_{\alpha=1}}{\left.\frac{\partial u(\alpha x)}{\partial \alpha}\right|_{\alpha=1}} = -\frac{u''(\alpha x)x^2\big|_{\alpha=1}}{u'(\alpha x)x\big|_{\alpha=1}} = -\frac{xu''(x)}{u'(x)}$$

This gives us the idea of the relative risk aversion.

2.2 Common Utility Functions

We have two measures of risk aversion, and it will be useful to think about how these measures change as x changes. You can think of this similarly as how does your risk aversion change as your wealth changes. It can either increase, decrease, or stay the same. So this gives us 6 types of utility functions:

Absolute Risk Aversion:

- *Increasing Absolute Risk Aversion (IARA):* $I_A(x,u)$ is increasing in x
  - Ex: Quadratic utility. $u(x) = x - \alpha x^2$. Then $u'(x) = 1 - 2\alpha x$ and $u''(x) = -2\alpha$. Therefore $I_R(x,u) = \frac{2\alpha}{1-2\alpha x}$
- *Decreasing Absolute Risk Aversion (DARA):* $I_A(x,u)$ is decreasing in x
  - Ex: Log utility. $u(x) = \log x$. Then $u'(x) = \frac{1}{x}$ and $u''(x) = -\frac{1}{x^2}$. Therefore $I_A(x,u) = \frac{1}{x}$
- *Constant Absolute Risk Aversion (CARA):* $I_A(x,u)$ is independent of x
  - Ex: Exponential utility. $u(x) = 1 - e^{-\alpha x}$. Then $u'(x) = \alpha e^{-\alpha x}$ and $u''(x) = -\alpha^2 e^{-\alpha x}$. Therefore $I_A(x,u) = \alpha$

Relative Risk Aversion:

- *Increasing Relative Risk Aversion (IRRA):* $I_R(x,u)$ is increasing in x
  - Ex: Exponential utility with $\alpha > 0$ (risk averse agent). Therefore $I_R(x,u) = \alpha x > 0$
- *Decreasing Relative Risk Aversion (DRRA):* $I_R(x,u)$ is decreasing in x
  - Ex: Exponential utility with $\alpha < 0$ (risk loving agent). Therefore $I_R(x,u) = \alpha x < 0$
- *Constant Relative Risk Aversion (CRRA):* $I_R(x,u)$ is independent of x

– Ex: Power utility. $u(x) = \frac{x^{1-\gamma}-1}{1-\gamma}$ for $\gamma \neq 1$ and $\log x$ for $\gamma = 1$. Then $u'(x) = x^{-\gamma}$ and $u''(x) = -\gamma x^{-\gamma-1}$. Therefore $I_R(x,u) = \gamma$

All of these utility functions can actually be summarized in one functional form called the *hyperbolic absolute risk aversion* or HARA. This has the form:

$$u(x) = \frac{1-\gamma}{\gamma}\left(\frac{ax}{1-\gamma}+b\right)^{\gamma}$$

If we compute the Arrow-Pratt measure:

$$\begin{aligned}
u'(x) &= \frac{(1-\gamma)}{\gamma}\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-1}\frac{a}{1-\gamma} = a\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-1} \\
u''(x) &= a(\gamma-1)\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-2}\frac{a}{1-\gamma} = -a^2\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-2} \\
\frac{u''(x)}{u'(x)} &= -\frac{-a^2\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-2}}{a\left(\frac{ax}{1-\gamma}+b\right)^{\gamma-1}} = a\left(\frac{ax}{1-\gamma}+b\right)^{-1} = \frac{a(1-\gamma)}{ax+(1-\gamma)b} = \frac{1}{\frac{a}{a(1-\gamma)}x+\frac{b}{a}}
\end{aligned}$$

In class, we let $\beta = \frac{a}{a(1-\gamma)}$ and $\alpha = \frac{b}{a}$ so we expressed the HARA as $I_A(x) = \frac{1}{\alpha+\beta x}$. If we let $\alpha = \frac{1}{c}$ and $\beta = 0$, then $I_A(x) = c$ which is CARA. If we let $\alpha = 0$ and $\beta = \frac{1}{c}$, then $I_A(x) = \frac{c}{x}$ which is DARA and therefore CRRA too ($I_R(x) = xI_A(x) = c$).[5]

2.3 Comparing Individuals

Our measures of risk aversion basically said: how does an agent's risk aversion change as we change their wealth? But this is keeping the agent (and therefore their utility function) the same. We also want to make comparisons between people - how can we say Person A is 'more risk averse' than Person B (and what does that even mean?)

For this part, let's consider two agents with preferences $\succsim$ and $\succsim'$ with Bernoulli utility functions $u(\cdot)$ and $v(\cdot)$, respectively.

Comparing Risk Aversion
The following are equivalent ways to "$\succsim$ is more risk averse than $\succsim'$":

1. $F \succsim \delta_x \implies F \succsim' \delta_x$, for all lotteries F and $x \in \mathbb{R}_+$
2. $c(F,u) \leq c(F,v)$, for all lotteries F
3. $\rho(F,u) \geq \rho(F,v)$, for all lotteries F
4. $I_A(x,u) \geq I_A(x,v)$, for all $x \in \mathbb{R}_+$
5. $I_R(x,u) \geq I_R(x,v)$, for all $x \in \mathbb{R}_+$
6. There exists a strictly increasing and concave function $\phi(\cdot)$ such that $u(x) = \phi(v(x))$

The first three definitions basically all say that the more risk averse agent is willing to accept a smaller amount of money to avoid the lottery. The last three definitions are saying that the risk averse agent

[5] I abbreviate increasing absolute/relative risk aversion as IARA/IRRA, which is what you'll commonly see. In class, we had IARA/IRRA refer to the *index* of absolute/relative risk aversion. Just another potentially confusing thing to watch out for

has a more concave utility function. Both of these make intuitive sense, so let's go through the math to see why they are equivalent. Here's what we're going to try and show:

$$\begin{tikzcd}
\boxed{(1)} \arrow[r, Rightarrow] & \boxed{(2)} \arrow[r, Leftrightarrow] \arrow[dl, Rightarrow] & \boxed{(3)} \\
\boxed{(6)} \arrow[u, Rightarrow] \arrow[r, Leftrightarrow] & \boxed{(4)} \arrow[r, Leftrightarrow] & \boxed{(5)}
\end{tikzcd}$$

[**1** $\implies$ **2**]

- Let $x = c(F, u)$, so $F \sim \delta_{c(F,u)} \implies F \succsim \delta_{c(F,u)}$. By (1), $F \succsim' \delta_{c(F,u)}$
- Using similar logic, $\delta_{c(F,v)} \sim' F$. By transitivity, $\delta_{c(F,v)} \sim' F \succsim' \delta_{c(F,u)}$
- Since $\delta_{c(F,v)} \succsim' \delta_{c(F,u)}$, then $v(c(F, v)) \geq v(c(F, u))$
- Since $v(\cdot)$ is increasing, then $c(F, v) \geq c(F, u)$

[**2** $\iff$ **3**]

- $c(F, u) \leq c(F, v) \iff \mu_F - c(F, u) \geq \mu_F - c(F, v) \iff \rho(F, u) \geq \rho(F, v)$ (since $\mu_F \geq 0$)

[**2** $\implies$ **6**]

- We can always find a strictly increasing function ϕ
  - By assumption, v is strictly increasing, so there exists an inverse v^{-1} that is strictly increasing
  - Define $\phi(x) = u(v^{-1}(x))$. Since u and v^{-1} are both strictly increasing, then ϕ must be strictly increasing too
- ϕ is also concave. Proof by contradiction: suppose ϕ is strictly increasing but not concave
  - Since ϕ not concave, there must exist $v, v' \in \mathbb{R}$ and some $\lambda \in (0, 1)$ such that

    $$\phi(\lambda v + (1 - \lambda)v') < \lambda\phi(v) + (1 - \lambda)\phi(v')$$

  - Find prizes $x, y \in X$ such that $v(x) = v$ and $v(y) = v'$ (v is continuous so this is always possible)
  - Define a lottery $F = \lambda\delta_x + (1 - \lambda)\delta_y$ (you can win either win x with probability λ or y with probability $1 - \lambda$)
  - The expected utility of this lottery under $\succsim'$ is $V(F) = \lambda v(x) + (1 - \lambda)v(y)$ (similarly for $U(F)$). Apply the $\phi(\cdot)$ transformation on $V(F)$:

    $$\phi(V(F)) = \phi(\lambda v(x) + (1 - \lambda)v(y)) < \lambda\phi(v(x)) + (1 - \lambda)\phi(v(y)) = \lambda u(x) + (1 - \lambda)u(y) = U(F)$$

  - Since $V(F) = v(c(F, v))$ and $U(F) = u(c(F, u))$, we can write the inequality as: $\phi(V(F)) < U(F) \implies \phi(v(c(F, v)) < u(c(F, u)) \implies u(c(F, v)) < u(c(F, u))$
  - Since $u(\cdot)$ is strictly increasing, then $c(F, v) < c(F, u)$. But this violates our starting assumption in (2)

[**6** $\implies$ **1**]

- $F \succsim \delta_x$ implies that $\int u(x)dF(x) \geq u(x)$
- Since ϕ is strictly increasing, its inverse ϕ^-1 must exist and is strictly increasing and convex
  - ϕ^{-1} is strictly increasing: $\int u(x)dF(x) \geq u(x) \implies \phi^{-1}\left(\int u(x)dF(x)\right) \geq \phi^{-1}(u(x)) = v(x) = V(\delta_x)$
  - ϕ^{-1} is convex: $\phi^{-1}\left(\int u(x)dF(x)\right) \leq \int \phi^{-1}(u(x))dF(x) = \int v(x)dF(x) = V(F)$ (by Jensen's inequality)
- Putting this together: $V(\delta_x) \leq \int u(x)dF(x) \leq V(F) \implies F \succsim' \delta_x$

[**6** $\iff$ **4**]

- Set up derivatives:
  - $u'(x) = \frac{d\phi(v(x))}{dx} = \phi'(v(x))v'(x)$
  - $u''(x) = \phi''(v(x))v'(x)v'(x) + \phi'(v(x))v''(x)$
- Use Arrow-Pratt measure formula
  - $I_A(x,u) = -\frac{u''(x)}{u'(x)} = -\frac{\phi''(v(x))v'(x)}{\phi'(v(x))} - \frac{v'(x)}{v''(x)} = -\frac{\phi''(v(x))v'(x)}{\phi'(v(x))} + I_A(x,v)$
- [$\implies$]
  - $v(\cdot)$ and $u(\cdot) = \phi(v(\cdot))$ are strictly increasing: $v' > 0$ and $\phi' > 0$
  - $\phi(\cdot)$ is concave: $\phi'' < 0$
  - $\therefore$ The signs in the first term are $-\frac{[-][+]}{[+]}$, which means it is positive. So $r_A(x,u) \geq r_A(x,v)$
- [$\impliedby$]
  - $r_A(x,u) \geq r_A(x,v) \implies$ First term is positive
  - Since $v(\cdot)$ and $u(\cdot) = \phi(v(\cdot))$ are strictly increasing, we would need $\phi'' < 0$ to make the first term positive. This means that $\phi(\cdot)$ is concave

[**4** $\iff$ **5**]

- $I_A(x,u) \geq I_A(x,v) \iff -\frac{u''(x)}{u'(x)} \geq -\frac{v''(x)}{v'(x)} \iff -x\frac{u''(x)}{u'(x)} \geq -x\frac{v''(x)}{v'(x)} \iff I_R(x,u) \geq I_R(x,v)$ (since $x \geq 0$)

This is all we need to have the equivalence. In class we did it slightly different (we didn't use (1)), and so we also had the proof for (6) $\implies$ (2). I'll include it here too:

[**6** $\implies$ **2**]

- By the definition of risk premium and certainty equivalent: $v(\mu_F - \rho(F,v)) = v(c(F,v)) = V(F)$. Apply $\phi(\cdot)$ transformation to both sides
- By (6): $\phi(v(\mu_F - \rho(F,v))) = u(\mu_F - \rho(F,v))$
- By vNM form and concavity of ϕ: $\phi(V(F)) = \phi\left(\int v(x)dF(x)\right) \geq \int \phi(v(x))dF(x) = \int u(x)dF(x)$
- Putting this together: $u(\mu_F - \rho(F,v)) \geq \int u(x)dF(x) = U(F) = u(c(F,u))$
- Since $u(\cdot)$ is strictly increasing and $u(\mu_F - \rho(F,u)) = u(c(F,u))$, the only way for this to hold is if $\mu_F - \rho(F,v) \geq \mu_F - \rho(F,u)$, which means that $\rho(F,u) \geq \rho(F,v)$

3 Application: Portfolio Composition

Suppose you had wealth w that you wanted to invest. You could either invest it in a safe asset that will give you a return of R next period or a risky asset where the return is some random variable $\tilde{x}$ that comes from a CDF $F_{\tilde{x}}(x)$, with support $[a, b]$. Let's call the share of your wealth that you invest in the risky asset as $\alpha \in [0, 1]$. What is the optimal α?

To get this, all you need to is to maximize your expected utility. Note that if I invest \$1 in the safe asset, I will get back R. So if I invest $(1-\alpha)w$, I will get back $(1-\alpha)wR$. So my total return from investments (for a given $\tilde{x} = x$) will be $\alpha w \cdot x + (1-\alpha)w \cdot R$. My expected utility $U(\cdot)$ will simply be my utility over this value, for all $\tilde{x} \sim F_{\tilde{x}}(x)$. The optimization problem is then:

$$\begin{aligned} \max_{\alpha} U(\alpha) &= E\left[u(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R)\right] \\ &= \int_a^b u(\alpha w \cdot x + (1-\alpha)w \cdot R) dF_{\tilde{x}}(x) \end{aligned}$$

As usual, we'll take a first order condition. Since our integrating interval does not depend α, we have a special case of Leibniz's rule and can just 'switch' the expectation and derivative:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} E\left[u(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R)\right] \\ 0 &= E\left[\frac{\partial}{\partial \alpha} u(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R)\right] \\ 0 &= E\left[u'(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R) \cdot w(\tilde{x} - R)\right] \end{aligned}$$

The question we then ask: is it ever optimal for the agent to invest completely in the safe asset (i.e. $\alpha^* = 0$)? Suppose that it is optimal, then the FOC becomes:

$$\begin{aligned} 0 = U'(\alpha^*) &= E\left[u'(wR) \cdot w(\tilde{x} - R)\right] \\ 0 = U'(\alpha^*) &= u'(wR)\left(E[\tilde{x}] - R\right) \end{aligned}$$

We will assume that u is a strictly increasing and concave function. Note that this means that $U(\alpha)$ is a concave function:

$$\begin{aligned} U(\lambda\alpha + (1-\lambda)\beta) =& E\left[u\big((\lambda\alpha + (1-\lambda)\beta)w \cdot \tilde{x} + (1 - (\lambda\alpha + (1-\lambda)\beta))w \cdot R\big)\right] \\ =& E\left[u\big(\lambda[\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R] + (1-\lambda)[\beta w \cdot \tilde{x} + (1-\beta)w \cdot R]\big)\right] \\ \geq& E\left[\lambda u\big(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R\big) + (1-\lambda)u\big(\beta w \cdot \tilde{x} + (1-\beta)w \cdot R\big)\right] \\ =& \lambda E\left[u\big(\alpha w \cdot \tilde{x} + (1-\alpha)w \cdot R\big)\right] + (1-\lambda)\lambda E\left[u\big(\beta w \cdot \tilde{x} + (1-\beta)w \cdot R\big)\right] \\ =& \lambda U(\alpha) + (1-\lambda)U(\beta) \end{aligned}$$

The inequality comes from the fact that $u(\cdot)$ is concave. Since U is concave, then we know that U' is a decreasing function. In particular, it means that we will not have a solution if $U'(0) < 0$ (since we need our FOC to be $U'(\alpha^*) = 0$). Now we can consider two cases:

1. $E[\tilde{x}] \leq R$: Since $u'(wR) > 0$ (u is strictly increasing), then $U'(0) \leq 0$. Therefore, the solution here is $\alpha^* = 0$ (the closest we could get to $U'(\alpha^*) = 0$ is by having $\alpha^* = 0$)
2. $E[\tilde{x}] > R$: Since $u'(wR) > 0$, then $U'(0) > 0$ and is optimal to invest $\alpha^* > 0$

The idea here is that it is only optimal to fully invest in the safe asset if $E[\tilde{x}] \leq R$ (the expected return from the safe asset is at least as good as the risky asset). But if $E[\tilde{x}] > R$, then it is optimal for the agent to invest even a little bit of their wealth in the risky asset. Note that this holds regardless of the degree of the agent's risk aversion (we just needed to assume concavity of u).

4 Comparing Risk

So far, we've looked at how different agents can view the same lottery in different ways. But preferences are all quite subjective - how can we *objectively* tell whether a lottery is "risky". To do this, we are going to introduce the concept of stochastic dominance. Throughout, we will be comparing two lotteries or CDFs F and G.

4.1 First Order Stochastic Dominance

The first way to compare lotteries is to say that one is objectively 'better' than the other one. We think of this by saying F is objectively 'better' than G if *all* (reasonable) expected utility maximizers would prefer F over G. By reasonable, we just mean that they prefer more money to less. This gives us the idea of first order stochastic dominance (FOSD):

First Order Stochastic Dominance
F first order stochastically dominates G (denoted by $F \geq_{FOSD} G$) if for all increasing Bernoulli utility functions $u(\cdot)$:

$$U(F) \geq U(G)$$
$$\iff \int_{x \in X} u(x)dF(x) \geq \int_{x \in X} u(x)dG(x)$$

So what kind of lotteries F and G would we need so that a risk averse, risk neutral, and risk loving agent would all agree on the preferred lottery? While these agents differ in their risk preferences, they all like more money to less money (u is assumed to increasing). So as long as larger prizes are more likely under F than under G, then all agents would agree that F is preferred to G. This gives us another way of expressing FOSD:

First Order Stochastic Dominance
$F \geq_{FOSD} G$ if and only if:

$$F(x) \leq G(x), \forall x \in X$$

Let's express this in a different way. Fix some prize x, and let $X_F \sim F(x)$, $X_G \sim G(x)$, then FOSD means that $P(X_F \leq x) \leq P(X_G \leq x)$. So under F you are less likely to get a prize worse than x as compared to G. That must mean that larger prizes are more likely under F, making F the better choice. Intuitively, these two ways of expressing FOSD seem to make sense but let's see the proof of why they are equivalent:

Proof: $F \geq_{FOSD} G \implies F(x) \leq G(x), \forall x \in X$

- $F \geq_{FOSD} G$ means that $U(F) \geq U(G)$ for any increasing $u(\cdot)$

- Let's define the support for the lotteries as $X = [a, b]$.[6] For some $\gamma \in [a, b]$, define $u(x) = \mathbb{1}\{x > \gamma\}$ (this is an indicator function). So $u(x) = 1$ if $x > \gamma$ and $u(x) = 0$ if $x \leq \gamma$.
- $u(x)$ is increasing, and so an agent with this utility must prefer F over G:

$$\begin{aligned}
\int_a^b u(x)dF(x) &\geq \int_a^b u(x)dG(x) \\
\int_a^\gamma 0 \cdot dF(x) + \int_\gamma^b 1 \cdot dF(x) &\geq \int_a^\gamma 0 \cdot dG(x) + \int_\gamma^b 1 \cdot \cdot dG(x) \\
0 + (1 - F(\gamma)) &\geq 0 + (1 - G(\gamma)) \\
F(\gamma) &\leq G(\gamma)
\end{aligned}$$

- Note that we are using the fact that $\int_\gamma^b dF(x)$ represents the total probability of getting x between γ and b (the max prize). So this is simply $P(X > \gamma) = 1 - P(X \leq \gamma) = 1 - F(\gamma)$
- Our choice of γ was arbitrary, so we have $F(\gamma) \leq G(\gamma), \forall \gamma \in X$

Proof: $F(x) \leq G(x), \forall x \in X \implies F \geq_{FOSD} G$

- Take any increasing utility function $u(\cdot)$. We need to show $U(F) \geq U(G)$ which is just the same as showing $U(F) - U(G) \geq 0$

$$\begin{aligned}
&\int_a^b u(x)dF(x) - u(x)dG(x) \\
&= \int_a^b u(x)(dF(x) - dG(x)) \\
&= u(x) \cdot (F(x) - G(x))\Big|_a^b - \int_a^b (F(x) - G(x)) \cdot du(x) \\
&= u(b)(\underbrace{F(b)}_{=1} - \underbrace{G(b)}_{=1}) - u(a)(\underbrace{F(a)}_{=0} - \underbrace{G(a)}_{=0}) - \int_a^b (F(x) - G(x)) \cdot \underbrace{du(x)}_{=u'(x)dx} \\
&= -\int_a^b \underbrace{(F(x) - G(x))}_{\substack{\leq 0 \\ \text{(by assumption)}}} \cdot \underbrace{u'(x)}_{\substack{\geq 0 \\ \text{(increasing)}}} dx \geq 0
\end{aligned}$$

- In the third line, we are using integration by parts, for which the general formula is: $\int_a^b u \cdot dv = u \cdot v|_a^b - \int_a^b v \cdot du$. Here, we have $u = u(x)$ and $dv = dF(x) - dG(x)$
- The inequality comes from the two negative terms canceling: $[-]\int[-][+]dx = [+]$

A few final remarks on this. First, you should note that $\geq_{FOSD}$ is a partial order. This means that it is not complete: you could take two lotteries and you may not be able to say that one FOSD the other one. Second, it quite easy to see that $\mu_F \geq \mu_G$ (the mean of F is larger) because $\int x dF(x) \geq \int x dG(x)$ (imagine you had a utility maximizer with $u(x) = x$). However this relation does not go the other way! A larger mean does not imply FOSD.

[6] If the support of F is $[a_1, b_1]$ and the support of G is $[a_2, b_2]$, then set $a = \min\{a_1, a_2\}$ and $b = \max\{b_1, b_2\}$

4.2 Second Order Stochastic Dominance

FOSD allowed us to find a lottery that was objectively better. The next question would be to find a lottery that is objectively riskier. Under FOSD, we wanted to find the lottery that all expected utility maximizers preferred. Now, we're going to just restrict ourselves to *risk averse* expected utility maximizers. Given lotteries F and G, if all risk averse people prefer F over G, then it must be because F is 'less risky' than G. This gives us the idea of second order stochastic dominance (SOSD).

In this part, we're going to assume that F and G have the same mean, i.e. $\mu_F = \mu_G$. That way we shut down the possibility that F is mechanically 'less risky' because it gives a higher return. The only reason F is being chosen over G is because of the dispersion in its prizes (which is how we naturally want to interpret riskiness).

> **Second Order Stochastic Dominance**
> F second order stochastically dominates G (denoted by $F \geq_{SOSD} G$) if for all increasing and concave Bernoulli utility functions $u(\cdot)$:
>
> $$U(F) \geq U(G)$$
> $$\iff \int_{x\in X} u(x)dF(x) \geq \int_{x\in X} u(x)dG(x)$$

Note that the key difference between the FOSD and SOSD is that we are adding the assumption of a concave utility function, which, we already know, is equivalent to risk aversion. Clearly if all expected utility maximizers prefer F over G, then the subset of risk averse agents must also prefer F over G. So we know that FOSD implies SOSD. Similarly, SOSD is a partial order too, and we can find an equivalent definition of it: (again, let's have $X = [a, b]$)

> **Second Order Stochastic Dominance**
> $F \geq_{SOSD} G$ if and only if:
>
> $$\int_a^x F(t)dt \leq \int_a^x G(t)dt, \forall x \in X$$

Graphically, this is saying that the area under the CDF F between a and x is always smaller than than the area under the CDF G between a and x, for any $x \in X$. Note that this is quite different to FOSD: there we just compared two points $F(x)$ and $G(x)$, but here we are comparing areas. Let's do the proof for the equivalence and then finish off with another way of explaining SOSD.

Proof: $F \geq_{SOSD} G \implies \int_a^x F(t)dt \leq \int_a^x G(t)dt, \forall x \in X$

- For some $\gamma \in [a, b]$, define a utility function as before $u(x) = (x - \gamma)\mathbb{1}\{x \leq \gamma\}$. This means that $u(x) = x - \gamma$ if $x \leq \gamma$ and 0 if $x > \gamma$
- $u(x)$ is increasing and concave (draw the picture and this will be easy to see). This means that $U(F) \geq U(G)$:

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$$
$$\int_a^\gamma (x-\gamma)\cdot dF(x) + \int_\gamma^b 0 \cdot dF(x) \geq \int_a^\gamma (x-\gamma)\cdot dG(x) + \int_\gamma^b 0 \cdot dG(x)$$

$$
\begin{aligned}
(x-\gamma)F(x)\Big|_a^\gamma - \int_a^\gamma d(x-\gamma)F(x) + 0 &\geq (x-\gamma)G(x)\Big|_a^\gamma - \int_a^\gamma d(x-\gamma)G(x) + 0 \text{ [integration by parts]}\\
(\gamma-\gamma)F(\gamma) - (a-\gamma)F(a) - \int_a^\gamma (1dx)F(x) &\geq (\gamma-\gamma)G(\gamma) - (a-\gamma)G(a) - \int_a^\gamma (1dx)G(x)\\
0\cdot F(\gamma) - (a-\gamma)\cdot 0 - \int_a^\gamma F(x)dx &\geq 0\cdot G(\gamma) - (a-\gamma)\cdot 0 - \int_a^\gamma G(x)dx\\
-\int_a^\gamma F(x)dx &\geq -\int_a^\gamma G(x)dx
\end{aligned}
$$

- Our choice of γ was arbitrary, so we have $\int_a^\gamma F(x)dx \leq \int_a^\gamma G(x)dx, \forall \gamma \in X$

Proof: $\int_a^x F(t)dt \leq \int_a^x G(t)dt, \forall x \in X \implies F \geq_{FOSD} G$

- First note that the assumption implies that $\int_a^x (F(t)-G(t))dt \leq 0, \forall x \in X$. In particular for $x=a$, this value is 0 (we are integrating over nothing) and similarly if $x=b$, then this value is 0 too because the means are the same (explanation below)
  - For $X \sim F(x)$, we can express the expectation as:
  $$E[X] = \mu_F = \int_0^\infty (1-F(x))dx - \int_{-\infty}^0 F(x)dx$$
  - Since F and G have finite and positive support $[a,b]$: $\mu_F = \int_a^b (1-F(x))dx$ (and similarly for G). Therefore $\int_a^b (F(t)-G(t))dt = \int_a^b (1-G(t)) - (1-F(t))dt = \mu_G - \mu_F = 0$ because they have the same means
- Take any increasing and concave utility function $u(\cdot)$. We need to show $U(F) \geq U(G)$ which is just the same as showing $U(F) - U(G) \geq 0$

$$
\begin{aligned}
&\int_a^b u(x)dF(x) - u(x)dG(x)\\
&= \int_a^b u(x)(dF(x) - dG(x))\\
&= u(x)\cdot(F(x)-G(x))\Big|_a^b - \int_a^b (F(x)-G(x))\cdot du(x)\\
&= -\int_a^b (F(x)-G(x))\cdot u'(x)dx\\
&= -\left(\int_a^x (F(t)-G(t))dt\cdot u'(x)\Big|_a^b - \int_a^b \left(\int_a^x (F(t)-G(t))dt\right)\cdot u''(x)dx\right)\\
&= -\underbrace{\left(\int_a^b (F(t)-G(t))dt\right)}_{=0} u'(x) + \underbrace{\left(\int_a^a (F(t)-G(t))dt\right)}_{=0} u'(x) + \int_a^b \left(\int_a^x (F(t)-G(t))dt\right) u''(x)dx\\
&= \int_a^b \underbrace{\left(\int_a^x (F(t)-G(t))dt\right)}_{\substack{\leq 0\\ \text{(by assumption)}}} \underbrace{u''(x)}_{\substack{\leq 0\\ \text{(by concavity)}}} dx \geq 0
\end{aligned}
$$

- To clarify, the third and fourth lines are just repeating what we did in the FOSD proof. In the next line, we do integration by parts again setting $u = u'(x)$ and $dv = (F(x) - G(x))dx$.

The last point in this section is to bring another interpretation of SOSD. Imagine you have a lottery $\tilde{x} \sim F$. Now we're going to define a new lottery $\tilde{y} \sim G$ that is a compound lottery: after you win x from F you play another lottery $\tilde{\varepsilon}$ where $E[\tilde{\varepsilon}] = 0$. So we can call your final prize from G as $y = x + \varepsilon$. The mean of G is clearly the same as the mean of F, but we've added another layer of risk. So in this sense G is 'riskier' than F (without improving the expected outcome), and so $F \geq_{SOSD} G$.

Let's make this a bit more general. I've made an assumption that $\tilde{\varepsilon}$ and $\tilde{x}$ are independent, but actually all we need is $E[\tilde{\varepsilon}|\tilde{x} = x] = 0$. In other words, conditional that you've received prize x from F, this second stage lottery should be mean zero. We call this concept a *mean-preserving spread.*

$$\tilde{y} \text{ is a mean-preserving spread of } \tilde{x} \text{ if } \tilde{y} = \tilde{x} + \tilde{\varepsilon}, \text{ where } E[\tilde{\varepsilon}|\tilde{x} = x] = 0$$

Note that this means that $E[\tilde{y}|\tilde{x} = x] = x + E[\tilde{\varepsilon}|\tilde{x} = x] = x$. The key point to take away here (without proof) is that $F \geq_{SOSD} G$ **if and only if** G is a mean preserving spread of F.