Formal Models: Section 2*

Zachary Lorico Hertz

26 January 2026

How to read this document

As noted in an earlier announcement, our scheduled section time is considerably shorter than it should be. This is problematic because we have two primary goals in section: to review content from the lecture and to apply it as practice for your assessments. As a result, we are offloading some of the former into these written handouts. As you read, you may feel familiar with or even mastery over some of the content, either because it landed well in lecture or because you picked it up easily from the reading.[1] If this is the case, by all means breeze through it here. I recommend, however, that you spend time ruminating on the material in this handout that feels unfamiliar or unintuitive to you. Take notes in the margins; underline; prove the math to yourself; jot down questions. These are meant to supplement the lecture and prepare for section and thus are best read between the two, while you may choose to read it before the lecture you should **definitely** read it before section, since we will proceed under the assumption that you have.

Review: from last week

Last section, we devoted most of our section time (both written and in-person) to getting you up to speed with the **mathematical language** upon which we will rely for the rest of the semester. Using our new alphabet, you hopefully now have the basic tools to formally represent choice sets and outcomes in the decision problems and games that we will analyze this semester.

We then defined **preference relations** as special types of binary relations that must satisfy two key conditions: they must be *complete* (which is to say, you must be able to compare any two elements in the set) and *transitive* (if you relate a to b and relate b to c, then a must relate to c). We noted the difference between *weak preferences* represented with $\succsim$ and *strict preferences* represented with $\succ$ is the possibility of *indifference* (denoted $\sim$). These properties help us avoid logical inconsistencies and ensure that our problems are tractable and solvable.

*Some parts of the discussion on useful properties draws on previous materials from Tak-Huen Chau, who himself drew on materials from David Foster and Jieun Kim. All writing (including its poor sense of humor and any errant mistakes) are mine.

[1] I have somewhat of a realist's view of the average graduate student's propensity to do readings in methods classes, however, which serves as a third motivation behind these handouts as a more-concentrated distillation of the content for you to more quickly consume. It's like an academic Soylent (or Huel, for those of you named Tak-Huen Chau).

Finally, we defined **lotteries**, which are probability distributions over outcomes, and their **expected utility**, which is simply the probability-weighted average of the utilities of the lottery's outcomes. We also noted that the *shapes* of actors' utility functions reflect their attitudes towards risk: those who are **risk-neutral** have utility functions with a linear form $a + bx$; those who are **risk-averse** have utility functions that are concave (for which $u''(x) < 0$); and those who are **risk-tolerant** have utility functions that are convex (for which $u''(x) \geq 0$).

A reminder on rationality

To quote the old adage, repetition breeds familiarity. So, if you'll pardon the potentially redundant reminder, we have a very particular definition of rationality in this class, that is in a manner of speaking quite simple. **Rational actors choose the action that maximizes their utility.**

Of course, there is a trick here. Actors choose a singular action from among a larger set of possible actions. Yet, we define rationality in terms of the actor's preferences over a set of *outcomes.* There is a difference, in other words, between the *consequences* that we map actors' preferences over and the *choices* they must make to realize these consequences. Often, these are one and the same. If, in the morning, I stop at my favorite coffee shop on the way to campus and decide to order a drink, the choice set and its consequences are relatively simple. I can order an iced coffee, hot coffee, or one discrete item out of a small set of menu items. The consequences are simple, too: if I order an iced coffee, I receive an iced coffee. Comparing utility is easy for simple finite action sets.

But some actions are *continuous, infinite sets.* When I pay for the coffee, I have the option to tip. I can tip one cent, one hundred dollars, or any of the $9,999$ other one-cent increments in between. To determine how I personally feel about any of these outcomes in relation to one another becomes complicated very quickly. **This is the purpose of a utility function**: it is a mathematical way for us to consider actors' preferences over continuous and infinite sets of outcomes.

Theorem 0.1. *If an action set A is **finite** and the preference relation $\succeq$ is **complete** and **transitive**, there exists a function u that maps outcomes onto the real number space (ie. $u : O \rightarrow \mathbb{R}$) such that for any pair of outcomes $x, y \in O$, $u(x) \geq u(y)$ if and only if $x \succeq y$.*

This is the basis of the rest of this course: **in each problem we are faced with, there should be a mathematical solution through which we can determine the action through which actors maximize their utility.** Actors should be aware of their sets of possible actions and outcomes, how their actions relate to outcomes, and what their preferences are over outcomes; and then choose the action that will realize their most-preferred outcome.[2]

Risk is not just a board game

Remember that we use *lotteries* to represent uncertainty, and *expected utility* to derive preferences under uncertainty. Here, I want to take a moment to drive in what utility functions are actually meant to represent.

[2] Note that this avoids substantive claims about the *content* of these preferences; emotional/extrinsic motivations still exist and in fact are compatible with this world!! Instead, this assumes that these potential motivations are factored into the decision-maker's understanding of how their actions connect to outcomes.

Recall from last section that a **lottery** is just a probability distribution over outcomes. If some outcome $x_i \in X$ occurs with probability p_i, we can write the lottery as $p = (p_1, p_2, \dots, p_n)$ where p_1 is the probability that x_1 occurs, and $\sum_i p_i = 1$. The lottery has an **expected value**, which is just the sum of each outcome x_i multiplied by its probability of occurring p_i, or $\sum_{i=1}^{n} p_i \times x_i$.

Note that the lottery's **expected value** differs subtly from the **expected utility**. The former tells us what we should expect to be the outcome of the lottery; the latter tells us how much a particular actor will value the epected outcome. Determining the **expected utility** of a lottery is rather straightforward and should appear familiar at this point. Similar to taking any probability-weighted average, we take the utility some actor receives from outcome x_i, which we define as $u(x_i)$ and multiply it by the probability of that event x_i actually occurring, which we define as p_i. Repeating this process for all n individual outcomes and taking their sum gives us the expected utility:

$$U(p) = \sum_{i=1}^{n} p_i u(x_i)$$

Earlier, I said that the utility function $u(x_i)$ was a way to represent preferences and rank outcomes. But, there is another purpose: utility functions tell us how actors feel about the risk presented by the uncertainty. This is demonstrated by the *shape* of the utility function.

Risk attitudes through utility functions

To illustrate this, let's start by considering a simple lottery. In this lottery, the probability that actors receive \$100 is $\frac{3}{4}$, and the probability that they receive \$0 is $\frac{1}{4}$. From the above, we can calculate the **expected monetary value** for this lottery, $\sum_{i=1}^{n} p_i \times x_i$, as $\frac{3}{4} \times \$100 + \frac{1}{4} \times \$0 = \$75$.

We consider three actors, all participating in this lottery, who also all have similar aggregate preferences: all three prefer having *more* money to having *less* money.[3] Where our actors differ, however, is in their *utility functions*, and thus these actors differ in how they feel about risk. I'll tell you up front that Person 1 is *risk-neutral*, Person 2 is *risk-averse*, and Person 3 is *risk-tolerant*, but see if you can follow the intuitions on why. We consider each person in turn:

Person 1: $u(x) = x + 4$

Remember that the utility function is a function of outcomes. It may help to think of it as an internal function, through which the actors convert outcomes to *"jollies"*[4] as a final measure of how satisfied they are with the final outcome.

Looking at the function $u(x) = x + 4$, we see that Person 1 has a utility function that is linear in form. We can calculate their expected utility from this lottery, using the general equation $U(p) = \sum_{i=1}^{n} p_i u(x_i)$, as follows. First, we note that by the above, the probability of winning \$100 is $p_{\$100} = \frac{3}{4}$ and the probability of winning nothing $p_{\$0} = \frac{1}{4}$. Second, we use the utility function to

[3] Who doesn't?

[4] Note that "jollies" are a term that are mostly used by Sean and less in favor with, well, everyone else. You'll usually see "utils" to refer to hypothetical units of utility or satisfaction. To keep our notes maximally ~synergetic~ with your lectures, I use "jollies" throughout, but I suggest using "utils" if you are ever in conversation with someone outside our department. An anonymous friend of mine referenced "jollies" when talking to an outside senior scholar who was … confused.

calculate how many jollies Person 1 receives from winning \$100: $u(100) = 100 + 4 = 104$ as well as how many jollies Person 1 receives from winning \$0: $u(0) = 0 + 4 = 4$. Then, the expected utility is trivial:

$$\begin{aligned}
U(p) &= \sum_{i=1}^{n} p_i u(x_i) && \text{(by definition)} \\
&= \frac{3}{4}(100+4) + \frac{1}{4}(0+4) && \text{(substituting in } p_{\$100}\text{, as well as } u(100) \text{ etc)} \\
&= 78 + 1 && \text{(arithmetic)} \\
&= 79
\end{aligned}$$

So, for our given lottery with an expected monetary value of \$75 we can see that Person 1 expects to receive 79 jollies. Consider, for a moment, that again this is a decision under *risk*, as defined last week. Person 1 can expect to receive \$75 in expectation, but is there a number of dollars that Person 1 would accept with certainty and value at least equally as much as the lottery that provides them with an uncertain \$75 in expectation?

Last week, we discussed the concept of *indifference*, which we can use to answer this question. We just proved that Person 1 gets 79 jollies from entering this lottery. Given that we also know that for any dollar amount x Person 1 gets a number of jollies $u(x)$ such that $u(x) = x + 4$, the question becomes whether we can find some certain amount x_{certain} such that $u(x_{\text{certain}}) = 79$. Hopefully you grasp the intuition: if Person 1 gets the same number of jollies from the certain payment as they get from the random lottery, they should be *indifferent*. So, we set this up mathematically:

$$\begin{aligned}
u(x_{\text{certain}}) &= 79 && \text{(from set up)} \\
x_{\text{certain}} + 4 &= 79 && \text{(because } u(x_{\text{certain}}) = x_{\text{certain}} + 4\text{)} \\
x_{\text{certain}} &= 75 && \text{(subtracting 4 from both sides)}
\end{aligned}$$

From this, we've found the certain payment that Person 1 values as much as the lottery: \$75. Does this look familiar to you? Hopefully it does: the expected monetary value of the lottery was also \$75. In short, we've just found that Person 1 is *indifferent* between entering the lottery and receiving a guaranteed \$75. Why is Person 1 indifferent between these two options? Person 1 is indifferent because the utility they can expect to receive from entering the lottery is equal to the utility they receive from receiving the lottery's expected monetary value, with certainty.

We refer to this as being **risk-neutral**: Person 1 is as happy taking on risk to obtain an expected \$75 as they are happy receiving it with certainty. It turns out that this property is true of all utility functions with a linear form, or those for which the second derivative is equal to 0. Is this true for everyone, full stop?

Person 2: $u(x) = \sqrt{x}$

We consider Person 2 now. Because it is the same lottery, the *expected monetary value* of the lottery will still be \$75, but Person 2's expected *utility* from this lottery differs. We still use the general equation $U(p) = \sum_{i=1}^{n} p_i u(x_i)$, and because it is the same lottery the probability of winning \$100 is still $p_{\$100} = \frac{3}{4}$ and the probability of winning nothing is still $p_{\$0} = \frac{1}{4}$.

But because the utility function is different, Person 2 receives a different amount of jollies if they win \$100; specifically, for Person 2 $u(100) = \sqrt{100} = 10$ and $u(0) = \sqrt{0} = 0$. As before, we use these to calculate Person 2's expected utility from the lottery:

$$
\begin{aligned}
U(p) &= \sum_{i=1}^{n} p_i u(x_i) && \text{(by definition)} \\
&= \frac{3}{4}(\sqrt{100}) + \frac{1}{4}(\sqrt{0}) && \text{(substituting in } p_{\$100}\text{, as well as } u(100) \text{ etc)} \\
&= 7.5 + 0 && \text{(arithmetic)} \\
&= 7.5
\end{aligned}
$$

So, for our given lottery with an expected monetary value of \$75 we can see that Person 2 expects to receive 7.5 jollies. Now, for Person 2, we also ask: is there a certain payment for which Person 2 is indifferent between the lottery and the certain payment? This will help us assess how they feel about risk. Again, we can start by defining this as some x_{certain} for Person 2 where $u(x_{\text{certain}}) = \mathbb{E}[u(\text{lottery})]$:

$$
\begin{aligned}
u(x_{\text{certain}}) &= 7.5 && \text{(from previous math)} \\
\sqrt{x_{\text{certain}}} &= 7.5 && \text{(because } u(x_{\text{certain}}) = \sqrt{x}\text{)} \\
x_{\text{certain}} &= 7.5^2 && \text{(square both sides)} \\
&= 56.25 && \text{(arithmetic)}
\end{aligned}
$$

So, we have found that Person 2 values a certain \$56.25 as much as they value the lottery with an expected monetary value of \$75. Restated, they are indifferent between entering the lottery and receiving a guaranteed \$56.25. Why would Person 2 be indifferent between taking home \$56.25 and entering the lottery if they expect to receive \$75 from the lottery?

We can try to intuit this by considering a case where we offered Person 2 a certain \$64, and seeing whether Person 2 prefers the certain \$64 to the lottery, prefers the lottery to the certain \$64, or is indifferent. Since $u(x) = \sqrt{x}$ for person 2, we note that $u(64) = \sqrt{64} = 8$ and the expected utility of the lottery for Person 2 was 7.5; because $8 > 7.5$ Person 2 prefers the certain \$64 to the uncertain expected \$75.

From this, we note that Person 2 is happier accepting less money with certainty than they would expect to obtain from the lottery. They are happy taking home \$56.25, rather than taking on an uncertain gamble. Restated, Person 2 is willing to give up expected value to avoid risk. We refer to this as being **risk-averse**. This is true of all utility functions that are concave: those where the function curves down or for which the second derivative is negative. But what about functions whose second derivative is positive (ie. the function is convex)?

Person 3: Risk Loving ($u(x) = x^2$)

For this, we consider Person 3. Hopefully by now you've gotten the hang of things, and can notice that for Person 3 $u(100) = 100^2 = 1000$ and $u(0) = 0^2 = 0$. Again, we use these to calculate Person 3's expected utility from the lottery:

$$
\begin{aligned}
U(p) &= \sum_{i=1}^{n} p_i u(x_i) && \text{(by definition)} \\
&= \frac{3}{4}(100^2) + \frac{1}{4}(0^2) && \text{(substituting in } p_{\$100}\text{, as well as } u(100) \text{ etc)} \\
&= 7500 + 0 && \text{(arithmetic)} \\
&= 7500
\end{aligned}
$$

So, for our given lottery with an expected monetary value of \$75 we can see that Person 2 expects to receive 7500 jollies. But how do they feel about risk? As we did previously, we try to find some x_{certain} for Person 3 where $u(x_{\text{certain}}) = \mathbb{E}[u(\text{lottery})]$:

$$
\begin{aligned}
u(x_{\text{certain}}) &= 7500 && \text{(from previous math)} \\
x_{\text{certain}}^2 &= 7500 && \text{(because } u(x_{\text{certain}}) = x^2\text{)} \\
x_{\text{certain}} &= \sqrt{7500} && \text{(take the square root of both sides)} \\
&= 86.60 && \text{(arithmetic, and rounding)}
\end{aligned}
$$

So, Person 3 is indifferent between taking home \$86.60 and entering the lottery with an expected monetary value of \$75. In other words, Person 3 needs to be offered *more* money with certainty before they prefer the certain money to the uncertain lottery. Person 3 is actually willing to *pay* to take on risk. We refer to this as being **risk-tolerant**.[5] This is true for all utility functions that are convex: those where the function curves up or for which the second derivative is positive and greater than 0.

A quick note on derivatives

Why all this focus on the second derivative? It won't come up much formally, but I want to include a quick explanation. You may recall from previous classes in calculus or physics that the second derivative gives us the rate at which a function's slope itself is changing; it is the "rate of change of the rate of change."

With utility functions, we refer to the first derivative (the rate of change) as *marginal utility*: how much does the actor benefit from a one-unit increase? Thus the second derivative gives us the rate of change of marginal utility.

Recall that for a **linear** utility function (risk neutrality), $u''(x) = 0$, so there's *constant* marginal utility. Each additional dollar is valued the same to the actor, regardless of whether they have made \$1 or \$100. For a **concave** function (risk aversion), however, $u''(x) < 0$ meaning the marginal utility of money is *decreasing.* This means that there are large increases in marginal utility at very low values of x, but smaller increases in marginal utility when x is already high. Restated, the risk-averse actor values gaining a dollar when they have made \$1 more than they value gaining a dollar when they have already made \$100. For a **convex** function (risk tolerance), $u''(x) > 0$, so marginal utility of money is *increasing.* This means that each additional dollar matters more than the last.

[5] In other work you might more commonly see this referred to as "risk-loving."

Restated, the risk-tolerant actor values gaining a dollar when they have made \$1 *less* than they value gaining a dollar when they have already made \$100.

Useful mathematical properties: distributions

I'm devoting the second half of this handout to review useful properties of distributions and derivatives. These properties are key to analyze the stochastic spatial voting model we will be working with in lecture and on your problem set. More generally, many of the models you will encounter in this class and political science writ large need some way to account for uncertainty. For example, voters might have idiosyncratic preferences, there might be measurement error, or there might be random shocks in the world. As analysts, we can model this uncertainty using **random variables** that follow particular **probability distributions**, whose properties we then exploit in our solutions.

What are probability distributions, and why do we care?

Reach your minds, if you will, back to math camp and you might recall that a **probability distribution** describes how probability is spread across different possible values of a random variable. For example, if we flip a fair coin, the distribution tells us that "heads" occurs with probability 0.5 and "tails" occurs with probability 0.5.

A coin flip is what we call a *discrete random variable*, because its outcomes are a countable, finite set (notice that we are beginning to build on the language we discussed last week!). For now, we are more interested in *continuous random variables*, or those with outcomes that are an infinite, uncountable set spanning some interval.

We describe distributions using two related functions:

- The **probability density function** (PDF), denoted $f(x)$, describes the relative likelihood of different values or outcomes. Essentially, the PDF tells us what values are more or less likely to occur, and higher values of $f(x)$ mean that values near x are more likely to occur. The PDF must satisfy $f(x) \geq 0$ for all x (since you cannot have a negative probability), and the total area under the curve must equal 1 (since probabilities sum to 1).
- The **cumulative distribution function** (CDF), denoted $F(x)$, tells us the probability that the random variable is less than or equal to x. Formally, $F(x) = P(X \leq x)$ where X is our random variable. The CDF is always between 0 and 1, and is non-decreasing.

These two functions are related: $\frac{dF(x)}{dx} = f(x)$. In other words, the probability density function is the derivative (or slope) of the cumulative distribution function.

Properties of symmetric, single-peaked distributions

In the stochastic voting model we present, we discuss a distribution with specific properties. I want to discuss what these properties mean and why they matter.

Suppose (on the problem set, perhaps) we give you a particular probability density function f that is:

- **Mean-zero**: This means that the expected value (average) of the random variable is 0. Formally, $\mathbb{E}[X] = 0$.
- **Symmetric around zero**: This means that the distribution looks the same on both sides of 0. Formally, $f(x) = f(-x)$ for all x.
- **Single-peaked**: This means that the PDF has one maximum point (the "peak"), and the PDF decreases as we move away from that peak in either direction.

Then, we also denote the corresponding cumulative distribution function as F. Recall from earlier that since the PDF is the derivative of the CDF, the CDF is the anti-derivative of the PDF (ie., you can take the integral of the PDF to find the CDF). These distributions have several useful properties:

Properties of the mean-zero, symmetric, single-peaked PDF:

- $f(x) > 0 \; \forall x \in \mathbb{R}$ (the density function is always positive)
- For $x < 0$, $\frac{df(x)}{dx} > 0$. Translated to English, for negative values of x, the derivative of the PDF is positive, or restated the density function is increasing as it approaches the mean from the left.
- For $x = 0$, $\frac{df(x)}{dx} = 0$. When $x = 0$, the density function is neither increasing nor decreasing. Its first derivative is equal to 0. This is the point of symmetry, where the peak of the distribution occurs.
- For $x > 0$, $\frac{df(x)}{dx} < 0$. The density function is decreasing as it moves away from the mean to the right.
- Putting this all together, we see that the PDF $f(x)$ is maximized when $x = 0$ and decreases as $|x|$ increases. Points closer to zero have higher density, meaning more probability mass is concentrated nearby.

Properties of the corresponding CDF:

- $F(x) \in [0, 1] \; \forall x \in \mathbb{R}$ (probabilities are always between 0 and 1)
- It is symmetric around zero. We can state this formally: $F(x) = 1 - F(-x)$
  - What does this mean? Recall from the definition of CDF that $F(x)$ gives the probability that our random variable is less than or equal to x. Similarly, $F(-x)$ gives the probability that our random variable is less than or equal to $-x$.
  - The equation $F(x) = 1 - F(-x)$ says that the probability of being below x equals the probability of being above $-x$.
  - This makes intuitive sense for a symmetric distribution: if the distribution looks the same on both sides of zero, then the probability of being below 2 should equal the probability of being above -2.
- $\lim_{x\to\infty} F(x) = 1$ (since the CDF maps the cumulative probability eventually, we "capture" all the probability)
- $\lim_{x\to-\infty} F(x) = 0$ (similarly, when we are very far to the left / very close to 0, almost no probability has accumulated)
- $\frac{dF(x)}{dx} = f(x)$ for all x (the CDF's slope equals the density; recall this was given earlier).

Useful mathematical properties: comparative statics and the chain rule

Comparative statics is a general term referring to analysis on how outcomes change when we change parameters in our model. For example, we might ask: "How does the probability that a voter chooses candidate D change when candidate R decreases a particular policy belief?"

We can structure this type of problem (analyzing how probabilities of some outcomes change with parameters) formally as $P = F(g(\theta))$ where:

- P is a probability we care about
- F is a cumulative distribution function
- $g(\theta)$ is some expression that depends on a parameter θ (e.g., a candidate's policy position or quality)
- θ is the parameter we're changing

Recall from calculus that a rate of change can be calculated with a simple derivative, so in this case we can calculate how our probability P changes in response to changes in θ by taking the first derivative of P with respect to θ. Specifically, we rely on the chain rule:

$$\frac{dP}{d\theta} = f(g(\theta)) \cdot \frac{dg(\theta)}{d\theta}$$

This says: the rate of change of the probability equals the density (evaluated at $g(\theta)$) times the rate of change of the threshold $g(\theta)$.

Summary

Moving on from *language* quickly into the material of this course, we have started to introduce the actual tools that we use in this class to model decision-making under uncertainty. This will be crucial to almost all models in political science, since there are countless sources of variation (random shocks, measurement error, idosyncratic preferences) that all might change how actors evaluate outcomes.

We spent more time examining how the shape of a utility function itself tells us how actors consider risk. Specifically, concavity can tell us whether an actor is indifferent to risk (risk-neutral), willing to pay to avoid it (risk-averse), or willing to pay to take it on (risk-tolerant). This will become clear when we get into games, since expected utilities are key to our models of behavior.

As a reminder, probability distributions (PDFs and CDFs) will be key to analyzing our models. In your future problem set, you will need to understand how specific properties (densities that are mean-zero, symmetric, and single-peaked) allow us to analyze the problem. Also important: when we ask (rhetorically or otherwise) how a probability changes in response to a parameter, we're asking about derivatives. This may require you to use the chain rule.

Concepts covered:

- Risk attitudes and utility function shapes
- Probability distributions (PDFs and CDFs)
- Properties of symmetric, single-peaked distributions
- Comparative statics with the chain rule