Case Study

When I lived in Los Angeles in the early 1990s I sometimes had to drive to Phoenix (AZ), a distance of about $D = 400$ miles along Interstate 10, on which the speed limit was 70 miles per hour (mph).

The faster I drove, the faster I got to Phoenix (good), but the greater the chance I got a speeding ticket (bad).

Evidently I needed to choose a compromise driving speed (not too slow, not too fast) – what’s the best possible compromise?

This is an example of a problem involving decision theory: I have to choose an action (here, this corresponds to picking a speed at which to drive) in the face of uncertainty (here, I don’t know whether or not I’ll get a ticket).

We talked about another decision-theory example on the first day of class: should a law regulating the dumping of refuse from ships into Monterey Bay be enacted or not, and if it’s enacted will this have a positive or negative effect on {the environment, the economy}?

There’s an established branch of statistics (and economics) devoted to studying how people can make optimal choices under uncertainty: decision theory.

One way to lay out the principles of this subject involves thinking about four ingredients:

- A set $A$ of available actions, one of which you will choose;
  - For each action $a$, a set $\mathcal{E}$ of uncertain outcomes describing what will happen if you choose action $a$;
- A set $\mathcal{C}$ of real-world consequences corresponding to the outcomes $\mathcal{E}$; and
- A utility function $U$ that quantifies your preferences for the consequences $\mathcal{C}$, with values of $U$ living on the number line and (without loss of generality) with large values of $U$ to be preferred.
Setting Up the Problem

Let's pretend that I drive at a constant rate $r$ and that I can achieve all possible speeds continuously between 70 and 90 mph.

Then $\mathcal{A}$ in this problem just consists of possible rates $r$ of driving speed in the interval $[r_{lo}, r_{hi}] = [70, 90]$, and $\mathcal{E}$ consists of pairs $[t = \frac{D}{r}, S(r)]$, where $t$ is the travel time and $S(r) = 1$ if I get a ticket going at rate $r$ and 0 if not.

$S(r)$ is like a random draw from a 0–1 population with $p(r)$ as the chance of getting a 1 (a speeding ticket) and $[1 - p(r)]$ as the chance of getting a 0 (no ticket).

Suppose that observational experience has shown me that the probability $p(r)$ of getting a ticket during the journey rises — roughly linearly — from 0 at $r_{lo} = 70$ mph to around $p_{hi} = 0.55$ at $r_{hi} = 90$ mph.

The hard part of applying decision theory turns out to be that all of the utility values have to be on the same scale, so that you can weigh the costs against the benefits of the various possible actions.

Let's say that speeding tickets cost $T = $150, and — if I get one — my yearly car insurance premium will go up by $I = $75.

Those are the costs of going too fast, so I also have to try to express the benefits of getting to my destination faster in monetary terms.

To quantify the advantage to me gained by decreasing the travel time, I discover after some thought that I would be willing to pay roughly $F = $100 per hour of reduction in driving time (I don't like long interstate drives).
Maximizing Expected Utility

As noted above, the utility function here has two parts: the gain from going faster, and the possible loss from getting a ticket.

At the slowest rate I'm contemplating it will take me \( \frac{400}{70} \) = 5.7 hours; at the fastest rate I'm considering the journey will take \( \frac{400}{90} \) = 4.4 hours; and in between the effective “monetary” gain to me will be

\[
F\left(\frac{D}{r_{lo}} - \frac{D}{r}\right) = 100 \left(\frac{400}{70} - 400\right). \quad (1)
\]

The monetary loss from the ticket would be
\( (T + I)S(r) = 225S(r) \), so the whole utility function is

\[
U(r) = F\left(\frac{D}{r_{lo}} - \frac{D}{r}\right) - (T + I)S(r)
= 100 \left(\frac{400}{70} - \frac{400}{r}\right) - 225S(r). \quad (2)
\]

Since big utility values are better than small ones, it seems like I should just find the value of \( r \) that maximizes utility, but I can't do that, because \( S(r) \) is random: I either get a ticket or I don't, and before I start driving I don't know which.

People have shown in this situation that the best you can do is to

Maximize the expected value of the utility function (or just maximize expected utility for short).

Here the only part of equation (2) that's random is \( S(r) \), which is either 1 with probability \( p(r) \) or 0 with probability \( [1 - p(r)] \).
Maximizing Expected Utility

Computing the expected value of $S(r)$ is like working out the mean of a population with $100p(r)$% 1s and $100[1 - p(r)]$% 0s:

$$E[S(r)] = p(r) \cdot 1 + [1 - p(r)] \cdot 0 = p(r).$$  \hspace{1cm} (3)

So the expected utility to be maximized is

$$E[U(r)] = \$F(D\frac{D}{r_{lo}} - \frac{D}{r}) - (T + I)p(r)$$

$$= \$100 \left( \frac{400}{70} - \frac{400}{r} \right) - \$225p(r).$$  \hspace{1cm} (4)

Now $p(r)$ is supposed to be linear, with the value 0 at $r = r_{lo} = 70$ and the value $p_{hi} = 0.55$ at $r = r_{hi} = 90$: this is just the straight line equation

$$p(r) = \frac{p_{hi}}{r_{hi} - r_{lo}}(r - r_{lo}) = 0.0275(r - 70).$$  \hspace{1cm} (5)

So finally we want to find the value $r^*$ of $r$ that maximizes

$$E[U(r)] = \$F(D\frac{D}{r_{lo}} - \frac{D}{r}) - (T + I)\frac{p_{hi}}{r_{hi} - r_{lo}}(r - r_{lo})$$

$$= \$100 \left( \frac{400}{70} - \frac{400}{r} \right) - \$6.1875(r - 70).$$  \hspace{1cm} (6)

This can be accomplished either

(a) by plotting $E[U(r)]$ against $r$ and reading off the graph the value $r^*$ that makes $E[U(r)]$ the biggest, or

(b) by calculus.
Maximizing Expected Utility

You can see that an $r$ of about 80 mph is best with this problem formulation; in fact (math interlude), the optimal $r^*$ is 80.4 mph — from the graph the global maximum of this function occurs at the only place where the first derivative is 0:

$$\frac{\partial}{\partial r} E[U(r)] = \frac{FD}{r^2} - \frac{(T + I) p_{hi}}{r_{hi} - r_{lo}}, \quad (7)$$

which when set to 0 yields the solution

$$r^* = \sqrt{\frac{FD(r_{hi} - r_{lo})}{p_{hi}(T + I)}}. \quad (8)$$

With the constants as given in the setup here the optimal speed is 80.4 mph, and at that speed my chance of a ticket is about 29%.
Sensitivity Analysis

![Graph showing expected utility vs. driving speed]

Notice, however, that the **biggest possible value** $U^*$ of $E[U(r)]$ is about $9.60, and if I wanted to have a **0% change** of getting a speeding ticket (by driving at $r = 70$ mph) the expected utility value from driving the speed limit is **$0**; in other words, I’m only **avoiding the loss of about $10** worth of time** while running a **substantial risk** of getting a ticket; in other words, this conclusion is **rather fragile**.

Another way to see this is to do a **sensitivity analysis**, varying some of the **constants** in the setup (which were only guesses, after all) to see how **stable** or **non-stable** the conclusion is.

**For example:**

- If I’m wrong about $F$ and the right value is **10% larger** than specified above, how much does $r^*$ change?
- If my estimate of $p_{hi}$ is **too low by 20%**, how much would that affect $r^*$?
Sensitivity Analysis (continued)

The optimal $r^*$ obtained above was

$$ r^* = \sqrt{\frac{FD(r_{hi} - r_{lo})}{p_{hi}(T + I)}}. $$

Because of the square root, increasing $F$ by 10% would increase $r^*$ by about 5%, and increasing $p_{hi}$ by 20% would decrease $r^*$ by about 10%.

For instance, with the constants as given except that $F$ goes from $100$ to $110$, $r^*$ would rise from 80.4 mph to 84.3 mph, and with the constants as given except that $p_{hi}$ goes from 0.55 to 0.66, $r^*$ would drop from 80.4 mph to 73.4 mph — you can see that the conclusion is fairly non-stable.

Another question that should always be asked is: How would you modify the basic problem formulation — what would you add to (or take away from) it — to make it more realistic?

Here are some ideas in this problem:

- The most important missing ingredient is that my chance of getting in an accident would also rise with $r$, and this would increase the cost of going faster.

- Speeding tickets are typically graduated in fee: 0–10 mph over the limit costs $T_1$, 10–20 mph over costs $T_2$, ...

- You should add (say) 20 minutes to the journey time to process the speeding ticket, which would act like a further penalty.

- The relationship between speed and ticket probability is almost certainly not linear; a bowl-shaped-up parabola having the value 0 at 70 mph would probably be more like it.

- I can’t really drive at a perfectly constant rate, and therefore the time it takes me is also random.
Other Examples

- Personally, on further reflection I'm not happy at having to suffer almost a 30% chance of getting a ticket at the optimal speed, so that means that I’ve over-valued the time I’ll save by going faster; and so on.

Here are two other decision-theory examples, both from the health sciences:

- How often should women get mammograms?

  The more often the better for finding breast cancer (benefit), but mammograms are not free, and there are risks of false positives (costs).

  Evidently the older a woman is, the more often she should be screened; is there an optimal age to start getting mammograms?

  People have used decision theory to arrive at the current recommendations: once a year starting at age 45–50, unless you have a family history of breast cancer (in which case you should start earlier) or you have one of the BRCA genes (maybe more often than once a year would be best).

- One way to measure the quality of health care in a hospital is to compare the observed mortality of its patients with the mortality you would have expected given how sick the hospital’s patients are when they’re admitted to the hospital.

  This requires a method for measuring patient sickness at admission.

  Typically there will be on the order of 100 variables in each patient’s medical record that are relevant to admission sickness.

  The more variables the better for making good predictions of who will live and who will die (benefit), but variables differ in how much they cost to collect data on — what’s the optimal subset of sickness variables?
Other Examples (continued)

With colleagues I’ve thought carefully about this problem:


We’ve used decision theory to show people how to choose subsets of sickness variables that achieve good cost-benefit trade-offs.

- A good book on decision theory in the health sciences is


- One last idea: experimental design and sample survey design are really decision problems — what’s the optimal (cost-benefit-tradeoff) data-gathering strategy, when you’re uncertain about how the data will come out?